Universidade de Brasília
Instituto de Ciências Humanas
Programa de Pós-Graduação em Filosofia

PROBABILIDADE E RACIOCÍNIO CIENTÍFICO – UM EXAME CRÍTICO DA ABORDAGEM BAYESIANA DE HOWSON & URBACH

BRENO HERMANN

Dissertação apresentada como requisito para a conclusão do curso de Mestrado em Filosofia da Universidade de Brasília, sob orientação do Prof. Dr. Agnaldo C. Portugal

Brasília, julho de 2013

AGRADECIMENTOS

Ao Agnaldo, que é o maior motivador deste trabalho. Em suas aulas foi que tive meu primeiro contato com o bayesianismo e que surgiu a ideia de realizar esta dissertação. Não fosse o estímulo, o diálogo, o apoio e a orientação segura, esta dissertação nunca teria vindo à luz. Em meio a tantos casos de mestrandos que enfrentam dificuldades para levar adiante suas dissertações, posso dizer que foi um grande privilégio e sorte ter podido trabalhar com o Agnaldo, por quem não posso deixar de registrar minha profunda gratidão, admiração e respeito.

Ao Gerson, amigo querido e sem dúvida um dos grandes professores do Departamento de Filosofia da Universidade de Brasília. Agradeço ao Gerson, com quem compartilho uma admiração profunda pela cultura germânica, por todo o conhecimento que me transmitiu em suas aulas sobre Hermenêutica e também fora delas, em conversas que sempre foram instigantes e agradáveis. Não fossem suas palavras e gestos de apoio na etapa final de redação, tudo teria sido bem mais difícil.

Ao Samuel, pelos comentários sempre pertinentes e precisos e que muito contribuíram para tornar este texto menos imperfeito.

À Marcia e ao Arthur, sem os quais minha vida seria estéril e sem sentido.

RESUMO

Este trabalho é uma investigação crítica da aplicação dos métodos probabilísticos para analisar o raciocínio científico, conforme o enfoque bayesiano proposto por Colin Howson e Peter Urbach no livro *Scientific Reasoning – the Bayesian approach.* Além da proposta dos dois autores, são apresentadas críticas ao bayesianismo, tanto aquelas respondidas por Howson e Urbach quanto as que não foram por eles consideradas. De modo a permitir uma melhor compreensão do bayesianismo, este é contrastado com a proposta teórica alternativa formulada por Deborah Mayo e conhecida como estatística do erro. O trabalho conclui pela importância do bayesianismo como instrumento de análise do raciocínio científico, sem prejuízo do fato de que algumas das críticas que lhe são dirigidas ainda carecem de respostas mais consistentes.

Palavras-chaves: filosofia da ciência, bayesianismo, Colin Howson, estatística do erro.

ABSTRACT

This dissertation is a critical investigation of the use of probabilistic methods to analyze scientific reasoning as proposed by Colin Howson and Peter Urbach in the book *Scientific Reasoning – the Bayesian approach.* In addition to the Bayesian approach of the two authors, both the criticisms addressed by them and the criticisms they have not taken into consideration are also presented. In order to allow a better understanding of the Bayesian approach, a contrast with the alternative theoretical approach elaborated by Deborah Mayo and known as error statistics is made. The dissertation contends that Bayesianism is still an important approach to investigate scientific reasoning, notwithstanding the fact that some criticisms still lack a coherent Bayesian response.

Keywords: philosophy of science, Bayesianism, Colin Howson, error statistics.

Sumário

INTRODUÇÃO

Esta dissertação pretende investigar criticamente a proposta bayesiana de Colin Howson e Peter Urbach para aplicar os métodos probabilísticos na análise do raciocínio científico. Para isso, pretende não apenas apresentar o enfoque bayesiano conforme elaborado por aqueles dois autores, como também examinar determinadas críticas formuladas ao longo das últimas décadas e, por fim, contrastá-lo com uma alternativa teórica específica, a chamada estatística do erro (*error statistics*). Pretende-se investigar de que maneira o bayesianismo se posiciona com relação à estatística clássica, quais críticas são levantadas por bayesianos contra os estatísticos clássicos e vice-versa, quais premissas são por eles compartilhadas e quais os separam de forma irreconciliável. Ao proceder com essa investigação, busca-se também identificar os pontos fortes e fracos do bayesianismo como forma de analisar o raciocínio científico. Com isso, serão levantados elementos que permitirão responder, em última instância, às seguintes questões: em que medida seria o bayesianismo útil para compreender o raciocínio científico, especialmente se contrastado à *error statistics*? Seriam válidas as limitações apontadas por um enfoque com relação ao outro? Nesse embate, qual deles se mostra mais rico de um ponto de vista heurístico?

Teorias científicas apresentam um caráter geral com relação às observações empíricas que devem explicar. As implicações da teoria vão, assim, além das informações contidas em qualquer conjunto limitado de observações empíricas. As leis de Mendel acerca da transmissão de características genéticas, por exemplo, pretendem-se válidas inclusive com relação a espécies da flora e da fauna extintas há séculos. Uma vez que é impossível atestar a aplicabilidade das leis de Mendel no caso dessas espécies, fica patente a amplitude que a teoria atribui a si própria, se comparada aos limites de qualquer conjunto de indícios colhidos a partir da realidade.

Identificada a questão acima, surge, porém, um novo problema. Uma vez que observações empíricas de escopo restrito constituem nosso único ponto de apoio seguro, como é possível acreditar no que dizem as teorias científicas?[1] Diante de um conjunto

[1] Ao longo de toda esta dissertação, os termos "teoria" e "hipótese" serão empregados de forma intercambiável. Embora esses dois termos não sejam sinônimos, preferi deixar de lado um exame detalhado das relações entre teorias e hipóteses, o que estaria fora do escopo do trabalho. Assim, ao discutir questões relativas a corroboração/confirmação e refutação/infirmação, empregarei os termos como se referindo praticamente à mesma coisa, o que não redundará em dano para os fins propostos. Agradeço ao Professor Samuel Simon Rodrigues por ter em alertado para essa questão e para a necessidade de deixar esse ponto claro para o leitor desde as primeiras páginas do trabalho.

limitado de experimentos e informações coletadas, como é possível crer, por exemplo, na validade das leis da gravidade, ou da teoria da relatividade, ou qualquer outro modelo que se pretenda de aplicação ampla? Tal questão, que foi analisada de forma pioneira e brilhante por Hume, consiste, precisamente, no chamado problema da indução.

Desde que formulado por Hume, o problema da indução passou a atrair a atenção de filósofos e cientistas, que se mostraram preocupados em dar-lhe uma solução.[2] Kant, por exemplo, fez uso de certos proposições *a priori* (por exemplo, "todo evento tem uma causa"), como forma de tentar justificar a validade da inferência de uma teoria a partir de um conjunto limitado de observações. Mas, como observa Howson, a questão da indução não se confunde com o problema da causalidade: o cerne da questão não é se todo evento tem uma causa, mas, antes, como determinar a(s) causa(s) específicas de um fenômeno determinado (Howson & Urbach, 2006, pg.1).

Outra tentativa de solucionar a questão da indução parte do chamado "Princípio da Uniformidade da Natureza". Tal princípio pode ser sintetizado, segundo Hume, da seguinte maneira: o futuro será sempre semelhante ao passado. A certeza de que o futuro repetirá o passado estaria, assim, na base das teorias científicas. Mas tal justificativa é claramente insuficiente (Howson & Urbach, 2006, pg.2). Em primeiro lugar, o princípio não diz em que aspectos o futuro repetirá o presente. Seu conteúdo é, portanto, vazio. Em segundo, o princípio teria de ser formulado de forma mais específica para justificar adequadamente a extrapolação das observações para a teoria. Ao aceitar a formulação de que "os metais se expandem quando aquecidos", tendo como base algumas observações e o princípio de que "o futuro repetirá o passado", estamos, na verdade, fugindo do problema que nos é colocado por Hume: como saber, a partir de um conjunto finito de observações sobre determinados metais, que todo e qualquer metal se expande quando aquecido? Em lugar de responder à questão de Hume, o chamado "Princípio da Uniformidade da Natureza" limita-se a varrer o problema para debaixo do tapete, sem fornecer qualquer base segura que permita ampliar as inferências derivadas dos indícios recolhidos.

Um dos aspectos perturbadores levantados pela questão da indução é que talvez não exista uma solução satisfatória para ela nos moldes que se deseja, isto é, que forneça ao raciocínio indutivo o mesmo grau de certeza da lógica dedutiva. Conforme se verá no curso da dissertação, ao elaborar sua visão da atividade científica baseada no

[2] Antes de Hume o problema da indução já havia chamado a atenção de filósofos medievais, notadamente mas não apenas Guilerme de Ockam, que se debruçaram sobre o assunto de uma perspectiva teológica.

teorema das probabilidades de Bayes, um dos objetivos de Howson e Urbach é, precisamente, demonstrar a existência de uma lógica indutiva particular, com princípios próprios e distintos dos princípios dedutivos e, dessa forma, garantir a consistência do raciocínio indutivo.[3] Até lá, permanecemos, porém, adstritos à posição de autores como Paul Fayerabend, que, diante da impossibilidade de superar o paradoxo lógico da indução, acreditam estar aberto o caminho para o ceticismo absoluto com relação à ciência.[4]

Nem todos, porém, se renderam às dificuldades apresentadas pelo problema da indução. A visão falsificacionista da ciência propugnada por Popper é uma tentativa de defender a racionalidade dedutiva da atividade científica contra o ceticismo, bem como de oferecer uma resposta às inquietudes levantadas pelo problema da indução. Para Popper, nenhuma teoria pode ser provada definitivamente por observações empíricas. É possível, porém, refutar teorias com base em observações. Além disso, as consequências dedutíveis de uma teoria podem ser verificadas pela observação. Nesse caso, Popper diz que a teoria foi corroborada, o que lhe confere certo valor epistêmico (Popper, 2002).

Sabe-se, hoje, que o fato de uma teoria ser corroborada tem um significado muito mais restrito do que se acredita tenha sido imaginado por Popper. Em virtude do fenômeno conhecido como "indeterminação das teorias científicas", há uma quantidade (presumivelmente) infinita de teorias que são confirmadas por um mesmo indício. A questão que se coloca, então, é menos a de que os indícios corroboram a teoria, do que saber como escolher entre teorias rivais igualmente adequadas aos indícios levantados. Popper não oferece solução para esse problema. A prática dos cientistas, por outro lado, mostra que o valor epistêmico de uma teoria é determinado, essencialmente, de forma comparativa. Eles hierarquizam teorias distintas corroboradas por um mesmo conjunto de indícios conforme o grau de credibilidade que atribuem a elas.

Howson aponta uma segunda crítica ao falsificacionismo popperiano, atribuída originalmente a Lakatos. Trata-se do fato de que, na maior parte das vezes, os indícios

---

[3] Embora Howson e Urbach não sejam claros a respeito, a leitura atenta de *Scientific Reasoning* indica que a preocupação dos autores é com a justificação do raciocínio indutivo nas ciências empíricas. Em nenhum momento os autores tentam aplicar o bayesianismo à matemática pura, que parece assim ficar de fora do escopo investigativo estabelecido. De fato, é difícil imaginar como o bayesianismo poderia ser aplicado em um contexto em que não se pode falar em indícios, nem em seu impacto sobre hipóteses matemáticas. Na verdade, a indução na matemática é um procedimento distinto da indução nas demais ciências, como a física. Na matemática, a indução, apesar de carregar esse nome, é um procedimento essencialmente dedutivo.

[4] Para ter uma visão mais aprofundada da visão "anarquista" de Feyerabend da Filosofia da Ciência, ver Feyerabend, Paul. **Against Method (quarta edição)**. Verso, maio de 2010.

que se obtêm por meio de experimentos científicos não são consequências dedutivas das teorias que, supostamente, confirmam. O exemplo utilizado por Lakatos refere-se às Leis de Newton, que dizem respeito às forças que agem entre objetos em geral. Predições específicas sobre o movimento dos planetas, por exemplo, só podem ser feitas pelas Leis de Newton com o auxílio de hipóteses auxiliares acerca da massa e posições dos planetas, a distribuição da massa no espaço, etc. Assim, as observações planetárias não constituem uma consequência lógica das Leis de Newton, muito embora se considere que elas corroboram essas mesmas leis (Howson & Urbach, 2006, pg. 4).

Quanto ao problema da indução, é bem sabido que Popper acreditava ser eminentemente dedutiva a lógica do raciocínio científico. Ele não estava sozinho em sua rejeição à indução. A mesma visão foi expressa pelo estatístico R. A. Fisher. Ambos combateram enfaticamente a opinião, prevalecente entre os séculos XVIII e XIX, de que a forma correta de avaliar hipóteses científicas, à luz de certo conjunto de indícios, é por meio do cálculo de probabilidades. No caso de Popper, a rejeição era motivada pela preocupação em negar que a base do conhecimento pudesse repousar sobre qualquer componente indutivo. No caso de Fisher, a preocupação principal era fugir do alegado subjetivismo que pode estar presente na avaliação de uma hipótese científica a partir da teoria das probabilidades (Howson, 2003, pg. 3).

O modelo falsificacionista de Popper padece de uma limitação adicional, relativa ao fato de que ele se aplica apenas a teorias determinísticas. Um primeiro ponto é, portanto, que esse modelo teria dificuldades em ser aplicado a teorias científicas probabilísticas como, por exemplo, as Leis de Mendel sobre a genética. Um segundo ponto é o seguinte: mesmo que concordemos com Popper que as teorias podem ser corroboradas por indícios (o que, em si mesmo já levanta um problema, dada a indeterminação das teorias científicas), a verdade é que tais indícios são obtidos por meio de instrumentos de medição que operam com margens de erro. Assim, mesmo no caso das teorias determinísticas, há situações em que o exame da compatibilidade lógica entre teoria e indícios envolve medidas experimentais que são necessariamente expressas em termos de probabilidades.

Popper e outros que partilham de sua visão falsificacionista da ciência não conseguem encontrar uma saída para um traço central da atividade dos cientistas: o fato de que, no dia a dia da prática laboratorial, para comparar hipóteses rivais, faz-se uso constante do raciocínio indutivo e de considerações probabilísticas. Reconhecido o fato inescapável de que as teorias científicas se estendem além de qualquer conjunto de indícios, bem como que inexiste uma forma de justificar tal extensão com o grau de certeza equivalente ao da lógica dedutiva, o "status" epistêmico conferido por cientistas a determinada hipótese é uma função da quantidade e qualidade dos indícios de que dispõem, podendo ser modificada por informações posteriores.

Essa característica central do método científico já foi reconhecida há tempos por filósofos da ciência e cientistas. É sabido que Poincaré, por exemplo, dizia que a única segurança do cientista de que as Leis de Newton não iriam falhar no próximo teste empírico é que tal falha seria "altamente improvável". O economista W.S. Jevons, por sua vez, escreveu que

> Apenas na medida em que uma indução particular se aproxima de uma indução perfeita é que ela se aproxima da certeza absoluta. A quantidade de incerteza corresponde à probabilidade de que outros objetos que os examinados possam existir e falsificar nossas inferências. A probabilidade de uma hipótese corresponde à quantidade de informação obtida por nosso exame. A teoria das probabilidades é o instrumento que nos impedirá de superestimar ou subestimar o conhecimento que temos.
>
> (Jevons, *apud* Howson & Urbach, 2006, pg. 6)

Surgia, então, a questão de como explicar e interpretar essas noções intuitivas, relativas à aplicação do raciocínio probabilístico ao método científico. Não foram poucas as tentativas empreendidas, que foram levadas a cabo por nomes tão ilustres quanto Rudolph Carnap, Harold Jeffreys, Bruno de Finetti, Frank Ramsay e E. T. Jaynes. Nesse "programa de pesquisa", há dois enfoques básicos. O primeiro entende serem objetivas, isto é, determinadas pela lógica apenas, as probabilidades atribuídas às teorias científicas. A atitude subjetiva do pesquisador com relação a determinada teoria é considerada irrelevante para o grau de probabilidade que se atribui. A premissa de fundo do enfoque objetivista é a de que seria possível justificar a indução apenas por meio da lógica. As dificuldades contra as quais esse enfoque se chocou foram, porém, insuperáveis, conforme se verá.

O segundo enfoque considera a probabilidade de uma teoria, essencialmente, como função da atitude do pesquisador com relação a ela. Nesse caso, a probabilidade específica da teoria é interpretada como função do "grau de crença" que o pesquisador tem com relação a ela. Chamada de probabilidade *subjetiva* ou *personalista*, esse enfoque está na base do programa metodológico bayesiano ou, simplesmente, do bayesianismo, nome que é devido ao famoso teorema que carrega o nome de seu criador, o Reverendo Thomas Bayes.

Em 1763, a Sociedade Real de Londres publicou um artigo de um de seus ex-membros, falecido dois anos antes. Tratava-se de “Um Ensaio com o propósito de resolver um problema na Doutrina das Probabilidades *(An Essay Towards Solving a Problem in the Doctrine of Chances)*, de autoria de um matemático e pregador não-conformista, o Reverendo Thomas Bayes. Nesse artigo, Bayes não chegou a dar tratamento matemático a suas ideias – tal formalização somente viria a acontecer no século XX –, tendo apenas as apresentado de forma intuitiva. Não obstante, tais ideias vieram a exercer posteriormente uma enorme influência. Uma grande variedade de teorias que carregam o nome de Bayes surgiu em áreas que vão da teoria das probabilidades a modelos de racionalidade científica e teorias de confirmação. Quais teriam sido as razões de tal encantamento?

Uma boa parte da atração do enfoque bayesiano pode ser atribuída a seu êxito em explicar princípios e práticas da metodologia científica a partir de um teorema não muito complicado, de premissas acerca do comportamento racional e de mudanças no grau de expectativas do investigador. Entre os fenômenos que o bayesianismo ajuda a compreender, podem-se mencionar a ocorrência de predições inesperadas; a preferência generalizada dos cientistas por hipóteses simples; a rejeição de hipóteses *ad hoc* na comparação de teorias rivais para explicar um mesmo conjunto de dados; a convicção de que conjuntos diversos de indícios constituem melhor evidência para determinada teoria de que um conjunto mais limitado; assim como o fato de que nem toda teoria é confirmada da mesma forma pelos indícios que dela decorrem.

Com base no teorema de Bayes, é possível oferecer explicações plausíveis para práticas da atividade científica em que o modelo popperiano se encaixa apenas com dificuldade. É possível, por exemplo, entender como uma instância de confirmação e outra de infirmação têm impactos diferenciados para uma hipótese submetida a teste. No caso de uma instância de confirmação *e*, há um efeito positivo sobre a teoria, na medida em que é aumentada a probabilidade posterior da hipótese *h*. Já no caso de uma instância de refutação de uma hipótese *h* cuja probabilidade prévia seja suficientemente alta, é possível creditar a ocorrência antes a hipóteses auxiliares do que à hipótese *h* propriamente dita. Nesse caso, o impacto negativo da refutação sobre *h* seria bem mais diluído do que o impacto positivo da confirmação (Portugal, 2004).

Um dos pontos fortes do teorema de Bayes está precisamente em conferir uma explicação capaz de relacionar todos esses elementos, os quais normalmente são considerados como desprovidos de qualquer vínculo entre si a partir de um enfoque hipotético-dedutivo. Além disso, o teorema de Bayes consegue solucionar de forma satisfatória questões que são enfrentadas com dificuldade por enfoques alternativos. É o caso, por exemplo, do paradoxo dos corvos.[5] Do ponto de vista estrito da lógica formal, a observação de um corvo preto confirma a hipótese que "todos os corvos são pretos" tanto quanto a observação de um objeto que não seja um corvo, nem preto. No caso de uma instância *e* de confirmação, o valor da confirmação tende a decrescer à medida que a probabilidade posterior da hipótese h aumenta. Se a hipótese h tiver probabilidade prévia já alta, o valor confirmatório da instância *e* será necessariamente baixo (Portugal, 2004). Nesse caso específico, o tratamento bayesiano resolve o problema ao explicar com clareza porque a observação de um corvo preto confirma com muito mais força a hipótese do que a observação de um objeto não-preto e não-corvo. Como a probabilidade de observação de um objeto não-preto e não-corvo é muito alta, seu valor de confirmação da hipótese será, necessariamente, baixo.

Os bayesianos entendem a atividade científica de forma bastante diversa do falsificacionismo popperiano. Os adeptos do teorema de Bayes acreditam que a razão pela qual a atividade científica está sempre a buscar novos indícios não é para falsear hipóteses, mas para diminuir a incerteza acerca de algum aspecto do mundo.

A formulação básica da teoria de Bayes é a seguinte:

$$P(A/B) = \frac{P(B/A)}{P(B)} \times P(A)$$

[5] O paradoxo dos corvos, também conhecido como paradoxo de Hempel, é um paradoxo lógico que surge ao se examinar a questão do que constitui um indício confirmatório de uma assertiva. Hempel (1965) descreve o paradoxo na forma de hipóteses: (1) Todos os corvos são pretos e (2) Tudo o que não é preto não é um corvo. Ambas as assertivas são logicamente equivalentes. Em todas as circunstâncias em que (2) é verdadeira, (1) também o é; em todas as circunstâncias em que (2) é falsa, (1) também é falsa. Isso estabelece uma condição de equivalência (*equivalence condition*, *EC*) entre as duas assertivas. Consequentemente, um indício que confirma (2), como a observação de um papagaio verde, também confirma (1). O fato de a observação de um papagaio verde confirmar a hipótese de que todos os corvos são pretos constitui o paradoxo, pois obviamente tal conclusão contraria a intuição. Uma possível solução para o paradoxo seria precisamente negar a condição de equivalência. Hempel, ao propor sua solução para o paradoxo, confirma, porém, a validade da condição de equivalência. Na verdade, Hempel soluciona o paradoxo ao negar que ele de fato existe. Para ele, o paradoxo apenas existe em função de que o analista dispõe de informações prévias sem as quais a observação de um objeto que não é negro, nem um corvo de fato confirmaria que todos os corvos são pretos. Tome-se a assertiva "o composto cloreto de sódio se torna amarelo ao ser objeto de combustão". Tal assertiva é equivalente a "um composto que ao ser objeto de combustão não se torna amarelo não é cloreto de sódio". Se efetuarmos um experimento com um composto desconhecido e, ao queimá-lo, verificarmos que ele não se torna amarelo, é natural concluir que não se trata de sódio e que, portanto, a hipótese de que o cloreto de sódio se torna amarelo ao queimar é confirmada. O paradoxo apenas surge se soubermos antecipadamente que tal composto é, por exemplo, cloreto de magnésio. Assim, para Hempel, o paradoxo apenas surge se o analista tiver informações prévias que tornam anti-intuitivo, embora verdadeiro, que o indício X, ao confirmar a assertiva Y, também confirma outras assertivas logicamente equivalentes a ela.

A fórmula significa que a probabilidade de o evento A ocorrer, tendo presente o evento B (termo à esquerda do sinal “=”), é calculada a partir da probabilidade de B dado A ou P(B/A), também chamada de verossimilhança de A; pela probabilidade prévia de B, ou P(B); e pela a probabilidade prévia de A, isto é, P(A). O propósito último da fórmula é avaliar a influência do evento B com relação a A, isto é, se, ocorrendo B, a probabilidade de A aumenta, diminui ou permanece inalterada.

É possível utilizar a fórmula de Bayes para avaliar a probabilidade de uma determinada teoria T ser verdadeira, tendo em vista o indício E, da seguinte forma:

$$P(T/E) = \frac{P(E/T)}{P(E)} \times P(T)$$

Nessa fórmula, afirma-se que a probabilidade da teoria T, dado indício E (termo à esquerda do sinal de igualdade), é igual ao quociente da expectativa subjetiva de E ocorrer, dada a teoria T, pela expectativa prévia de E, multiplicado pela probabilidade prévia – isto é, anterior à ocorrência de E – da teoria T.

A fórmula de Bayes levanta várias questões de interesse. P(E/T) parece ser o termo menos problemático da equação, uma vez que expressa o grau de expectativa da realização de E, dado T. Entre os axiomas da teoria das probabilidades está aquele segundo o qual qualquer probabilidade deve ser expressa, necessariamente, por um número entre 0 e 1, inclusive. No caso, 0 corresponde à probabilidade nula de o evento ocorrer (0% de probabilidade), enquanto 1 corresponde à probabilidade máxima (100%). Assim, se E for logicamente dedutível partir de T, P(E/T) terá o valor 1. Se, por outro lado, T for uma teoria estatística, então P(E/T) expressará a probabilidade de E, assumindo-se T como verdadeira (por exemplo, 0,6 ou 60%).

Atribuir valores a P(E) e P(T) é tarefa mais complexa. No caso de P(E), por exemplo, há várias dificuldades relativas ao que fazer no caso de E ser verdadeiro – isto é, ocorrer –, mas sua probabilidade não ser 1. Wesley Salmon indica, por exemplo, que para saber a probabilidade prévia de E seria necessário conhecer a probabilidade prévia de ~T (não T), o que por sua vez traz complicações adicionais (Salmon, 1998). Já com relação à probabilidade prévia de T, isto é, P(T), muitos bayesianos admitem não haver um valor objetivo. Eles assumem uma postura subjetivista e afirmam que P(T) nada mais é do que o grau de expectativa do investigador na ocorrência de T. Reconhecem, de acordo com essa postura, que pessoas diferentes podem atribuir valores distintos a P(T).

A interpretação subjetivista do teorema de Bayes parece ser, de fato, seu "calcanhar de Aquiles" na tentativa de nos ensinar algo com relação às pretensões de confirmação no raciocínio cientifico baseadas em inferências. Muitos críticos do bayesianismo indicaram ser equivocado atribuir valor à probabilidade prévia de uma teoria com base em expectativas individuais de verdade. Pioneiros do bayesianismo como Frank P. Ramsey, Bruno De Finetti e Leonard J. Savage tentaram contornar essa dificuldade. Para eles, determinada probabilidade nada mais é do que a expressão do quanto alguém estaria disposto a apostar na veracidade de uma proposição. Mas tais tentativas nem sempre se mostraram exitosas. Como saber se expectativas pessoais expressas probabilisticamente satisfazem os axiomas do cálculo de probabilidades, como exige o Teorema de Bayes? No mundo real, tais axiomas são violados todo o tempo. É possível, por exemplo, que os graus de expectativa de um indivíduo em Q e os graus de expectativa em ~Q (não Q) não somem 1, como requer um daqueles axiomas (Curd, M & Cover, J. A. 1998, p. 635).

Foi para responder a essa e outras objeções que os defensores do bayesianismo formularam o chamado argumento do caderno de apostas holandês *(the Dutch book argument)*. Jogadores profissionais dizem que um caderno de apostas holandês foi preparado contra alguém quando este aceita uma série de apostas tais que, independentemente do resultado de cada rodada isolada, é certo que irá ao final perder dinheiro. Nenhum ator racional apostaria dessa forma. É possível provar que, para evitar as consequências desastrosas do caderno de apostas holandês, basta que os graus de expectativa do apostador obedeçam aos axiomas da teoria das probabilidades. Uma conclusão possível é que quando os bayesianos subjetivistas interpretam probabilidades, eles estão trabalhando com certo grau de idealização. Eles tomam graus de expectativa pessoal como se o indivíduo fosse perfeitamente racional (Curd, M & Cover, J. A., 1998, p. 635). Tal idealização poderia, por outro lado, ser considerada como consequência do "caráter normativo" que muitas vezes o bayesianismo assume e que nem sempre é expresso de forma clara por seus defensores (Strevens, 2005).

As questões acima indicam a importância do bayesianismo como tentativa de interpretar o raciocínio científico. Mesmo aqueles que o rejeitam não se furtam a enfatizar que qualquer tentativa de compreender o método científico não pode prescindir do bayesianismo. Várias das questões deixadas em aberto por esse enfoque (as alegações de que não descreve corretamente a objetividade do raciocínio científico; de que a analogia entre situações de aposta e avaliação de teorias científicas não se aplica; ou, ainda, que não auxilia no processo de escolha se duas teorias distintas forem confirmadas pelo mesmo conjunto de indícios, isto é, em caso de indeterminação) constituem elementos importantes que têm sido explorados em programas de pesquisa em filosofia da ciência. Muito da influência desse enfoque deve ser creditada aos trabalhos de Colin Howson e Peter Urbach, especialmente ao livro *Scientific Reasoning: The Bayesian Approach,* que constituiu a primeira tentativa de fôlego de apresentar o bayesianismo de um ponto de vista filosófico.

Não obstante sua relevância, o bayesianismo é um assunto que parece ter sido objeto de pouco interesse na produção dos programas de pós-graduação nas universidades brasileiras. Tal fato já constituiria, por si só, justificativa de algum peso para que venha a ser tema de uma dissertação de mestrado, à luz da possibilidade de contribuir para o repertório acadêmico brasileiro em filosofia da ciência. Ao mesmo tempo, o reconhecimento gozado pelo enfoque de Howson e Urbach não impediu o surgimento de propostas alternativas que, ao rejeitar as premissas básicas do bayesianismo, recorrem a técnicas da estatística para interpretar o raciocínio científico de forma diversa.

Dentre os modelos alternativos, talvez o que tenha logrado maior repercussão tenha sido aquele que ficou conhecido como estatística do erro (*error statistics*), que tem em Deborah Mayo um de seus principais expoentes. De forma distinta ao bayesianismo, esse enfoque faz uso de técnicas estatísticas não para mensurar a probabilidade de hipóteses, mas para modelar padrões que são úteis para identificar, controlar e aprender a partir de erros experimentais. As diferenças fundamentais entre o tratamento bayesiano e o tratamento da *error statistics,* decorrem, em última instância, das diferenças entre as visões personalista e frequentista da teoria das probabilidades:

> As duas posições principais esposadas com relação à teoria das probabilidades resultam de uma tentativa de definir uma escala de probabilidades que possa ser empregada em consonância com os processos usuais do pensamento racional. Para uma escola, o grau de crença em uma proposição... constitui a noção básica com relação à qual a escala numérica deveria ser ajustada. A outra escola ressalta como, na vida quotidiana, o conhecimento das frequências relativas de ocorrência de um conjunto particular de eventos, no contexto de uma série de repetições, exerce influência na conduta (humana); ela sugere, assim, que é através da correlação com a frequência relativa que a medida numérica de probabilidade tem um significado mais direto para a mente humana.

(Pearson, *apud* Mayo, 1997)[6]

O foco de atenção da *error statistics* é encontrar métodos ou procedimentos adequados para testar hipóteses. Para isso, faz-se uso de métodos e modelos clássicos da estatística, como testes de significância (*significance tests*) e métodos de confiança de intervalos (*confidence intervals*). O uso de probabilidades prévias, por exemplo, é rejeitado pela estatística do erro, a menos que estejam baseadas em cálculos que possam ser justificados objetivamente, e não em avaliações subjetivas. A probabilidade é aqui utilizada para permitir avaliar hipóteses rivais a partir do aprendizado com erros identificados ao longo do procedimento científico. O cálculo de probabilidades não tem como objetivo direto ampliar o grau de confiança do pesquisador em sua hipótese.

Conforme afirmado no início desta introdução, este trabalho pretende investigar criticamente o bayesianismo e Howson & Urbach tendo, como contraponto, a *error statistics* de Deborah Mayo. O trabalho estará dividido em quatro capítulos. No primeiro capítulo da dissertação, o bayesianismo será apresentado de forma sistemática conforme o livro clássico de Howson e Urbach, bem como outros textos consagrados de autores bayesianos. Serão examinadas as soluções propostas por esse enfoque a pontos controversos da atividade científica, como o paradoxo dos corvos, a indeterminação das teorias científicas e as hipóteses *ad hoc.*

O capítulo seguinte será dedicado às críticas formuladas ao enfoque bayesiano, conforme foram analisadas e respondidas por Howson e Urbach nas três edições de *Scientific Reasoning* (1989, 1993 e 2006). No capítulo 3, serão apresentadas críticas que autores como Sober, Gillies e Deborah Mayo fizeram ao bayesianismo e que não foram objeto de resposta direta por Howson e Urbach.[7] O capítulo 4, por fim, servirá à apresentação da estatística do erro conforme formulada inicialmente em *Error and the*

---

[6] "The two main atitudes held to-day towards the theory of probability both result from an attempt to define the probability number scale so that it may readily be put in gear with common processes of rational thought. For one school, the degree of confidence in a proposition… provides the basic notion to which the numerical scale should be adjusted. The other school notes how in ordinary life a knowledge of the relative frequency of occurrence of a particular class of events in a series of repetitions has again and again an influence on conduct; it therefore suggests that it is through its link with relative frequency that a numerical probability measure has the most direct meaning for the human mind". (Pearson, E. S. On Questions Raised by the Combination of Tests Based on Discontinuous Distributions, citado em Mayo, Deborah, Error Statistics and Learning from Error: Making a Virtue of Necessity. In *Philosophy of Science,* vol. 64. Chicago: The University of Chicago Press, pp 195-212, 1997)

[7] Por uma questão de conveniência e porque Colin Howson foi o principal responsável pela resposta às críticas e o que mais publicou sobre o assunto dos dois coautores, frequentemente este trabalho mencionará apenas Howson e não Howson e Urbach.

Growth of Experimental Knowledge (1996). A dissertação é encerrada por uma conclusão em que se procura responder às perguntas formuladas anteriormente.

1. O ENFOQUE BAYESIANO DO RACIOCÍNIO CIENTÍFICO

Neste capítulo, será feita uma apresentação geral do bayesianismo como instrumento para a análise do raciocínio científico. Algumas questões mencionadas na introdução serão aprofundadas. Tanto quanto possível, pretendo concentrar-me nos argumentos filosóficos e evitar apresentações excessivamente formalizadas, de modo a não dificultar a fluência do texto. Conforme indicado, a apresentação do bayesianismo terá, como base, o livro de Howson e Urbach *Scientific Reasoning- the Bayesian approach*, cuja última edição é de 2006.

Da exposição que se segue será possível depreender, como assinalado por Strevens (2006), que o enfoque bayesiano é, em sua natureza, tanto descritivo quanto normativo. Ao longo da leitura de *Scientific Reasoning,* fica claro que o propósito de Howson e Urbach não é apenas esclarecer determinadas questões relativas ao raciocínio científico. Antes, eles estão igualmente preocupados em recomendar ao cientista procedimentos metodológicos que, no seu entender, são mais corretos para a atividade científica do que outros eventualmente empregados, notadamente o falsificacionismo de Popper.

CONFIRMAÇÃO BAYESIANA

Nesta seção inicial, Howson e Urbach têm em vista um duplo objetivo ao introduzirem o leitor no universo bayesiano. A consecução de tal objetivo exige que a argumentação seja desenvolvida etapas sucessivas. Em um primeiro momento, Howson e Urbach apresentam as noções básicas com que opera o bayesianismo, como confirmação e infirmação. Essa apresentação é acompanhada por exemplos em que o leitor toma conhecimento de que a aplicação do Teorema de Bayes forneceria explicações adequadas e iluminaria aspectos inusitados para fenômenos salientes da prática científica. Em um segundo instante, os autores dão um passo adicional: eles se propõem a mostrar que o bayesianismo é superior a enfoques alternativos, em especial ao falsificacionismo popperiano.

A premissa de base para a aplicação do Teorema de Bayes ao raciocínio científico é a constatação de que as informações obtidas por meio da observação exercem um papel essencial para confirmar ou infirmar uma teoria ou hipótese. No enfoque bayesiano, o aprendizado a partir da experiência é, essencialmente, uma questão quantitativa. Assim, poderíamos dizer que:

e confirma ou apoia *h* caso $P(h/e) > P(h)$;

e infirma *h* caso $P(h/e) < P(h)$;

e é neutro com relação a *h* caso $P(h/e) = P(h)$;

Tais asserções, que inicialmente podem parecer simples definições, constituem, na verdade, o núcleo do argumento bayesiano. De acordo com esse argumento, e conforme se verá nos exemplos das seções seguintes, para o bayesiano, noções preexistentes de confirmação e infirmação podem ser objeto de uma explicação satisfatória a partir do Teorema de Bayes.

O Teorema de Bayes, que foi apresentado na introdução, nos ensina que a probabilidade posterior de uma hipótese depende de três fatores: P(e/h), P(e) e P(h). Uma vez conhecidos esses fatores, é possível determinar se *e* confirma ou não *h*, bem como calcular P(h/e). Howson indica que, na prática, é comum que as probabilidades sejam conhecidas apenas de forma imprecisa, mas isso não inviabiliza a aplicação do Teorema de Bayes como forma de chegar a uma justificativa para a inferência científica. Para dadas probabilidades prévias P(e) e P(h), diz-se que *e* refuta *h* caso a probabilidade posterior de *h*, dado *e*, seja igual a zero, isto é, P(e/h) = 0. Nessa situação, o grau de infirmação é máximo. De forma semelhante, o grau máximo de confirmação ocorre, por sua vez, quando a probabilidade posterior de *h*, dado *e*, for igual a 1, isto é, P(e/h) = 1, o que se dá ao *e* decorrer logicamente de *h*. Ao estabelecer valores para P(e/h), estamos adstritos aos axiomas do cálculo de probabilidades, um dos quais estabelece que a probabilidade de um evento deve ser expressa por um número entre 0 e 1. Hipóteses probabilísticas admitem valores intermediários para P(e/h). Nesse caso, quanto maior for o valor, maior será a confirmação e vice-versa.

Na teoria subjetivista das probabilidades, que é a adotada por Howson e Urbach,[8] ao atribuir um valor para a probabilidade prévia P(h), o cientista é influenciado mais por fatores de índole pessoal ou psicológicos, do que por qualquer fato objetivo a respeito da proposição *h*. Dessa forma, a atribuição de determinado valor a P(h) expressa menos a probabilidade de veracidade da hipótese, do que o quão provável o cientista acredita que a hipótese seja. Uma interpretação possível para isso, já aludida, é a de Ramsey, para quem a probabilidade prévia deve ser interpretada como a "atitude"

[8] Conforme se verá adiante, determinados autores bayesianos que antecederam a Howson e Urbach, como Jaynes e Rosencrantz, acreditam que as probabilidades prévias, apesar de exibirem certo grau de subjetividade, podem e devem ter sua atribuição de valores constrangida por fatores objetivos. Howson e Urbach, porém, consideram que qualquer tentativa de constranger a liberdade do cientista ao atribuir probabilidades prévias é ilusória. Critérios supostamente objetivos seriam, para eles, arbitrários e de difícil justificação.

do cientista com relação à hipótese, a qual pode ser medida por meio do montante que estaria disposto a apostar em sua veracidade.

No dia a dia dos laboratórios, é sabido que determinado indício confirma uma hipótese h com especial força se tal indício for particularmente inesperado ou surpreendente. Tomem-se mais duas formulações do Teorema de Bayes além da primeira apresentada na introdução (Howson e Urbach, 2006, p. 97):

$$P(h/e) = \frac{P(e/h)}{P(e)} \times P(h) \ = \ \frac{1}{P(h) \ + P({\sim}h) \ \times \ \frac{P(e/{\sim}h)}{P(\frac{e}{h})}}$$

No caso de *e* ser uma consequência de *h*, em que P(e/h) = 1, o grau de confirmação de *h* é inversamente proporcional a P(e), conforme a segunda formulação, ou a P(e/~h), na terceira formulação. Assim, quanto menor for a probabilidade prévia P(e), ou seja, quão mais improvável for *e*, maior será a confirmação com relação a *h*. Da mesma forma, quanto menor for a probabilidade de *e* dado que *h* seja uma hipótese falsa, isto é, P(e/~h), maior será a probabilidade posterior de *h*. Na verdade, P(e) relaciona-se com as probabilidades de a hipótese *h* ser verdadeira ou falsa pela seguinte fórmula: P(e) = P(e/h)×P(h) + P(e/~h)×P(~h) (Howson & Urbach, 2006, p. 19). Note-se, por fim, que a probabilidade posterior da hipótese *h* depende da probabilidade prévia da própria hipótese, isto é, P(h).

Howson enfatiza que a relação inversa entre a probabilidade de determinado indício e seu poder de confirmação constituem uma consequência direta do Teorema de Bayes. Teorias que rejeitam a avaliação probabilística de hipóteses em razão de alegada “falta de objetividade” não conseguem explicar adequadamente o fenômeno. Fiel ao objetivo de exibir as virtudes do bayesianismo frente às insuficiências do popperianismo, Howson lembra que Popper tentou, em vão, formular uma explicação a partir de enfoque do cálculo de probabilidades que fosse menos subjetivo. De fato, Popper reconheceu que, para calcular o poder de corroboração de um indício *e* com relação a determinada hipótese *h*, deve-se partir dos valores de P(e/h) e P(e). Popper mediu o valor de confirmação ou corroboração, para empregar a terminologia bayesiana, que *e* atribui a *h* – tal valor corresponderia à diferença entre P(e/h) e P(e). Howson lembra, porém, que Popper nunca explicitou o que entendia pela probabilidade do indício *e*. É de se supor que Popper nunca admitiria qualquer conotação subjetivista para P(e), pois, caso o fizesse, isso representaria séria ameaça a seu enfoque para o raciocínio científico. Howson acredita que Popper deveria ter em mente uma noção lógica das probabilidades. Tal noção nunca chegou, porém, a ser desenvolvida. De qualquer forma, o ponto essencial é que o enfoque popperiano nunca logrou apresentar qualquer explicação para o benefício, do ponto de vista epistêmico, que um indício improvável aporta para a probabilidade de determinada hipótese *h* ser verdadeira.[9] Essa limitação seria, por outro lado, explicada de forma adequada pelo bayesianismo. (Howson & Urbach, 2006, p. 99; Popper, 2002, apêndice IX).

O bayesianismo é plenamente compatível com a possibilidade de uma teoria ser refutada pela experiência, conforme exige a concepção popperiana de ciência. Se uma hipótese *h* implica *e* como consequência, então P(e/h) = 1 e P(h/~e) = 0, desde que P(h)>0. No jargão bayesiano, isso significa que h é infirmada em grau máximo quando é refutada. Tal refutação persistirá caso novos indícios similares a *e* sejam coletadas. Se uma nova evidência *e´* for logicamente consistente com *e*, teremos que P(h/~e&e´) = 0. De forma análoga, o teorema também mostra que uma teoria é confirmada por toda evidência que seja dela decorrente.

Outra questão interessante decorrente da aplicação do Teorema de Bayes para compreender o raciocínio científico, a qual tampouco é explicada adequadamente por Popper, é a noção, também intuitiva, de que não faz sentido repetir indefinidamente determinada experiência com o objetivo de confirmar uma hipótese. É possível mostrar que, ao repetir infinitamente a mesma experiência, a partir de certo ponto a sequência de probabilidades posteriores chega a um limite. Isso significa que

$$\lim_{n\to\infty} P(h/e_1 \ldots \& e_n) = \lim_{n\to\infty} P(h/e_1 \ldots \& e_{n-1})$$

Ao estender o número de experimentos ao infinito, o limite das probabilidades posteriores tende a 1, o que mostra a falta de sentido de continuar a repetir a experiência indefinidamente (Howson & Urbach, 2006, p. 94). É verdade, porém, que o teorema não permite saber a partir de que ponto as repetições se tornam desnecessárias, pois isso

[9] Na primeira edição (1935), ainda em alemão, da *Logic of Scientific Discovery*, Popper analisou em detalhe a teoria frequentista das probabilidades. Segundo essa teoria, a probabilidade deve ser definida como a razão entre acontecimentos favoráveis e o total de acontecimentos possíveis. Em versões posteriores até 1994, essa análise foi complementada por material com maior grau de formalização relativo à Teoria das Probabilidades. No *Postscript* de 1982, por outro lado, Popper apresentou uma interpretação distinta para a noção de probabilidades, chamada de teoria da propensão. Segundo essa nova teoria, a probabilidade de um evento seria a propensão inerente às condições que produzem sequências de eventos.

exigiria conhecer a sequência de probabilidades posteriores que o cientista atribui a cada repetição, o que é impossível.

A confirmação bayesiana explica, ainda, as razões pelas quais uma versão restrita *h´* de uma teoria *h* é capaz de confirmar esta apenas até certo ponto. Tome-se, por exemplo, para confirmar a Teoria de Newton, uma versão simplificada que explique a queda livre dos corpos. Uma vez que a versão restrita *h´* é uma decorrência lógica da versão completa, então P(h) ≤ P(h´) (Howson & Urbach, 2006, p. 17). A partir da acumulação sucessiva de indícios decorrentes de *h´* (que também são decorrências de *h*), é possível deduzir que o limite da probabilidade posterior de *h* é dado pelo fator P(h)/P(h´).[10] Isso significa que a probabilidade posterior de *h* nunca poderá ser maior do que a razão P(h)/P(h´). Com isso, a probabilidade prévia de *h´*, P(h´), que aparece no denominador, funciona como um limite para a probabilidade posterior de *h*. O que isso significa do ponto de vista da prática científica? Significa que a versão mais reduzida *h´* confirma a hipótese original *h* apenas até certo ponto e o faze tanto menos quanto maior for o valor da probabilidade prévia P(h´). O fato de uma versão mais reduzida de uma teoria ou hipótese apenas confirmar esta até certo ponto constitui uma constatação comum no dia a dia dos cientistas. Trata-se, assim, de mais uma situação que o bayesianismo consegue explicar a contento.

As explicações descritas acima estão baseadas no fato de que a evidência acumulada pela repetição sucessiva de um mesmo experimento aumenta a probabilidade de novas repetições produzirem resultados idênticos (Howson & Urbach, 2006, p. 96), o que justificaria não continuar a partir de determinado ponto. Tal possibilidade é negada por Popper, que rejeita categoricamente a possibilidade da indução no raciocínio científico. Howson ressalta a impossibilidade de justificar a falta de sentido de repetir indefinidamente um mesmo experimento como forma de confirmar uma hipótese a partir do quadro descritivo-normativo de Popper, o que demonstra uma vez mais a superioridade do bayesianismo. Tentativas de continuadores de Popper, como

---

[10] Imagine-se uma série de predições dedutíveis de *h* e de *h´*. Uma vez que essas predições se realizem e, portanto, confirmem ambas as hipóteses, é possível calcular as probabilidades posteriores para ambas as teorias, de modo que $P(h/e_1 \wedge e_2 \wedge \ldots \wedge e_n) = \frac{P(h)}{P(e_1 \wedge e_2 \wedge \ldots \wedge e_n)}$ e $P(h'/e_1 \wedge e_2 \wedge \ldots \wedge e_n) = \frac{P(h')}{P(e_1 \wedge e_2 \wedge \ldots \wedge e_n)}$. Ao combinar essas duas equações e eliminar o denominador comum, tem-se $P(h/e_1 \wedge e_2 \wedge \ldots \wedge e_n) = \frac{P(h)}{P(h')} \times \frac{P(h')}{P(e_1 \wedge e_2 \wedge \ldots \wedge e_n)}$. Como o valor do último elemento à direita da equação não pode ser maior do que 1, isso significa que a probabilidade posterior de *h* nunca poderá ser maior que o fator $\frac{P(h)}{P(h')}$. Ver Howson e Urbach, 2006, p. 95.

Musgrave,[11] de elaborar tais justificativas caem no erro de implicitamente pressupor justamente aquilo que é negado por Popper, isto é, a indução.

Em resumo, a análise bayesiana se funda nas probabilidades prévias da hipótese h - P(h) - e do conjunto de indícios e – P(e) –, bem como na probabilidade de e ocorrer, uma vez que se tome h como verdadeira, isto é P(e/h). O bom funcionamento da "máquina bayesiana" (Strevens, 2005) depende do compromisso, do qual nenhum bayesiano foge, de atribuir valores definidos a essas probabilidades prévias, nem que seja por meio de suposições, as quais são limitadas somente pelo respeito aos axiomas do cálculo de probabilidades.

Nem sempre a atribuição de probabilidades prévias é tarefa simples. No caso de P(e/h), a situação é muito facilitada no caso de a própria hipótese h atribuir certa probabilidade física a e. É o caso, por exemplo, da probabilidade de se obter o resultado "cara" ou o resultado "coroa" ao jogar uma moeda não viciada. Intuitivamente, a probabilidade, tanto em um caso como em outro, é ½. Já para calcular a probabilidade prévia do indício e, é possível evitar a atribuição de um valor hipotético se o pesquisador tiver em mãos todas as informações que o permitam recorrer ao teorema da probabilidade total, segundo o qual

$$P(e) \;= P\left(\frac{e}{h_1}\right)P(h_1) \;+\; P\left(\frac{e}{h_2}\right)P(h_2) \;+\cdots$$

Para fazer uso do teorema, é necessário que h_1, h_2,... formem um conjunto de hipóteses mutuamente exclusivas, no sentido de que apenas uma delas é verdadeira. Assim, de posse dos valores de P(e/h_i) e das probabilidades prévias P(h_i) das hipóteses rivais, é possível chegar a um valor único e correto para a probabilidade subjetiva P(e).

Ao colocar em funcionamento a "máquina bayesiana", o cientista pode, porém, esbarrar em duas sérias dificuldades. É possível que alguma(s) das hipóteses rivais não atribuam um valor determinado para a probabilidade do indício e. Nesse caso, o analista precisaria atribuir um valor hipotético para P(e/h_i) – o que nos leva de volta ao problema inicial de subjetivismo - ou recorrer a hipóteses auxiliares que o permitam

[11] Em *Popper and Diminishing Returns from Repeated Tests* (*Australasian Journal of Philosophy*, volume 69, pp. 124-134), Alan Musgrave tenta salvar o argumento popperiano ao propor que, após determinado número de repetições, cujo número ele não especifica, o cientista estaria autorizado a inferir a generalização de que o experimento sempre produzirá um resultado similar. Tal generalização seria, então, considerada como conhecimento de fundo (*background knowledge*). O problema é, como nota Howson (2006, p. 97) que, ao dar esse passo, o cientista popperiano estaria tacitamente invocando a noção de indução que ele tradicionalmente rejeita.

estimar um valor para essa(s) probabilidade(s). Trata-se de uma dificuldade real, a qual não constitui um empecilho insuperável para bayesianos, que acreditam ser possível estimar probabilidades prévias com razoável veracidade, mesmo em situações aparentemente complexas.

A segunda dificuldade é que o teorema da probabilidade total pode não se aplicar se as hipóteses h_1, h_2,... não forem mutuamente exclusivas. Nessa situação, é possível conceber uma hipótese residual para o caso de nenhuma das demais hipóteses ser correta. Mesmo assim, as probabilidades P(e) e P(e/h_i) terão, muito provavelmente, de ser determinadas com base em estimativas, o que traz de volta, uma vez mais, a crítica do subjetivismo subjacente ao enfoque bayesiano.

Conforme se verá no próximo capítulo, essa e outras críticas contundentes foram objeto de reflexão pelos bayesianos, que elaboraram respostas com o intuito de preservar a integridade de seu modelo.

O PARADOXO DOS CORVOS

A esta altura, está claro que, para o bayesianismo, a confirmação é uma questão de grau. Isso fica muito evidente a partir da explicação que o bayesianismo ensaia para o chamado paradoxo dos corvos ou paradoxo da confirmação, proposto por Hempel, em 1945, no artigo *Studies in the Logic of Confirmation.* Hempel descreve o paradoxo por meio da hipótese:

(1) Todos os corvos são pretos;

Em termos lógicos, essa hipótese é equivalente à sua negação, isto é:

(2) Tudo o que não é preto não é um corvo.

Em todas as circunstâncias em que (2) é verdadeira, (1) também o é. De forma similar, em todas as circunstâncias em que a afirmativa (2) for falsa, a afirmativa (1) também será falsa. É evidente que a observação de um corvo preto confirma a hipótese (1), assim como sua equivalente (2). O paradoxo se verifica ao se estar diante de uma instância que confirma (2), como, por exemplo, quando se observa uma maçã verde. Pelo raciocínio anterior, a observação da maçã verde, ao confirmar (2), também confirma (1), o que é contraintuitivo.

Duas condições específicas são exigidas pelo paradoxo:

(1) Hipóteses do gênero “todos os corvos são pretos” são confirmadas pela observação de um objeto que seja, ao mesmo tempo, “corvo” e “preto” (*condição de Nicod*);
(2) Hipóteses equivalentes do ponto de vista lógico são confirmadas pelo mesmo indício (*condição de equivalência ou condição de Hempel*);

É possível tentar solucionar o paradoxo com base na premissa de que as duas condições não são mutuamente consistentes. Assim, uma das formas de resolvê-lo seria pelo abandono de uma das duas condições identificadas acima ou por sua relativização. Esse não foi, porém, o caminho seguido por Hempel, que preferiu manter a validade das duas condições e, com isso, aceitar o resultado de que a observação de uma maçã verde confirma a hipótese de que todos os corvos são negros. Na verdade, Hempel considera que o paradoxo é apenas aparente, em função de possuirmos informações prévias sem as quais não haveria qualquer desconforto no fato de uma maçã verde constituir um indício confirmador de que todos os corvos são pretos.

Para entender a solução que Hempel dá ao paradoxo, tome-se a hipótese “Todo e qualquer sal de sódio, ao ser queimado, fica amarelo”. Em seguida, Hempel pede ao leitor que considere o que acontece quando determinado cientista tem em suas mãos uma amostra de gelo que, ao ser queimada, obviamente não fica amarela:

> Esse resultado confirmaria a afirmativa “tudo o que, ao ser queimado, não se tornar amarelo, não é sal de sódio” e, consequentemente, em função da condição de equivalência, confirmaria também a hipótese original. Por que isso nos impressiona como paradoxal? A razão fica clara quando comparamos a situação prévia com um experimento em que um objeto, cuja constituição química é ainda desconhecida, é queimado e não se torna amarelo. Nessa segunda situação, ficamos sabendo, em uma análise subsequente, que o objeto não contém qualquer traço de sal de sódio. Esse resultado, nós deveríamos concordar, é o que se deveria esperar com base na hipótese original (...). Assim, a informação fornecida pelo experimento consiste em um indício confirmador da hipótese. Nos casos aparentemente paradoxais de confirmação, nós não estamos em realidade julgando as relações entre o indício e a hipótese (...). Tacitamente, nós introduzimos uma comparação da hipótese com um conjunto mais amplo de indícios, que consiste no indício observado em conjunção com uma quantidade adicional de informação que nós temos à nossa disposição. No caso específico, esse conjunto de indícios inclui a informação (1) de que a substância utilizada na experiência é gelo; e (2) de que o gelo não contém sal de sódio. Se assumimos que esse conjunto de informações adicional é dado, então claramente o resultado deixa de confirmar a hipótese original. Se, por outro lado, formos cuidadosos o suficiente para evitar essa referência tácita a informações adicionais, então o paradoxo desaparece
>
> (Hempel, 1945, pp.1-26)

Por que Howson e Urbach analisam o paradoxo dos corvos? Já se disse acima que os autores têm um propósito duplo: eles querem não apenas mostrar que o bayesianismo explica inúmeras questões complexas da filosofia da ciência, mas também que o faz melhor do que propostas teóricas alternativas. Dessa forma, o objetivo de Howson e Urbach é mostrar que a solução que o bayesianismo oferece para o paradoxo dos corvos é melhor do que a proposta por Hempel.[12]

Ao explicar o paradoxo dos corvos com base no Teorema de Bayes, Howson argumenta que se trata de uma situação menos problemática do que inicialmente se supõe. Assim como na solução proposta por Hempel, também na solução bayesiana são conservadas a condição de equivalência e a condição de Nicod, muito embora esta última não seja, para Howson e Urbach, um princípio de confirmação universalmente válido.

Seja CP a expressão de um objeto "corvo" e "preto" e ~C~P a expressão de um objeto que não é "corvo", nem "preto". Seja, ainda, Θ a expressão dos valores possíveis para expressar a proporção dos corvos que são pretos (nossa hipótese h diz, precisamente, que Θ=1). É possível mostrar (Howson & Urbach, 2006, p. 101) que:

$$\frac{P(h/CP)}{P(h)} = \frac{1}{\sum \theta P(\theta)} \ \& \ \frac{P(h/{\sim}C{\sim}P)}{P(h)} = \frac{1}{P({\sim}C{\sim}P)}$$

Segundo a primeira equação, a razão da probabilidade posterior pela probabilidade prévia de h é inversamente proporcional ao fator $\sum\theta P(\theta)$. Isso significa que, se a probabilidade de todos os corvos serem pretos for alta, a observação de um corvo preto terá um impacto pequeno do ponto de vista de confirmação da hipótese h. Se, por outro lado, a probabilidade de grande parte dos corvos não serem pretos for alta, o grau de confirmação será considerável.

A segunda equação, por sua vez, diz respeito à confirmação exercida pela observação de um objeto que não é corvo nem preto. Nesse caso, a razão da probabilidade posterior pela probabilidade prévia de h é inversamente proporcional ao fator P(~C~P). Nesse caso, como a probabilidade de observação de um objeto com essas

[12] Note-se que há várias soluções possíveis para o paradoxo, inclusive uma famosa apresentada por Quine em *Natural Kinds*, no livro *Ontological Relativity and Other Essays*, em que o autor argumenta que a solução está no reconhecimento de que determinados predicados, por ele chamados de "tipos naturais", se comportam de forma distinta de outros predicados do ponto de vista da indução. Mesmo no campo bayesiano há alternativas à solução de Howson e Urbach, apresentadas, por exemplo, por Earman, Eels, Gibson, Hosaisson-Lindenbaum, Hintikka, entre outros.

características é alta, o impacto sobre a probabilidade posterior será muito pequeno, o que confirma nossa intuição sobre o assunto. Assim, é possível compreender, por meio da explicação bayesiana, porque, intuitivamente, a observação de um corvo preto confirma muito mais a hipótese “todos os corvos são pretos” do que a observação de uma maçã verde. Na solução bayesiana para o paradoxo dos corvos, foram conservadas como válidas tanto a condição de equivalência quanto a condição de Nicod. A condição de equivalência deve, porém, ser atenuada: embora um indício *e* confirme tanto a hipótese original quanto outra logicamente equivalente a ela, o efeito de confirmação será diferenciado em um caso e noutro, o que é intuitivamente correto.

A solução proposta pelo bayesianismo para o paradoxo dos corvos foi, durante muito tempo, considerada um fator de peso em favor da tese de que esse enfoque teria logrado apresentar uma justificativa adequada para as inferências indutivas, ao fundamentá-las em um argumento *a priori*. Alguns analistas indicaram, porém, que não obstante o valor da solução bayesiana, essa seria uma conclusão apressada (Strevens, 2006). Tome-se uma vez mais a hipótese do paradoxo dos corvos, segundo a qual, ao constatar que todos os corvos observados até agora são pretos, a visão de um corvo preto confirma a hipótese de que todos os corvos são pretos. Tome-se, agora, o paradoxo proposto por Goodman: todos os corvos vistos até agora são pretos, os demais são verdes. O fato de que todos os corvos encontrados até este momento sejam pretos confirma tanto a hipótese original quanto a hipótese de Goodman. O fluxo de probabilidade corre, assim, não apenas em favor das hipóteses que estipulam que todos os corvos são pretos, como também em favor de outras opções possíveis, como a proposta por Goodman. Isso significa que a hipótese de que todos os corvos são pretos apenas é favorecida pelos indícios recolhidos no caso de a probabilidade a ela atribuída previamente ser mais alta do que a das hipóteses rivais. Se a probabilidade prévia atribuída não for mais alta do que a das rivais, o raciocínio bayesiano não será útil como instrumento para justificar a indução. É por essa razão que muitos bayesianos abandonaram a tentativa de considerar seu enfoque uma justificativa a priori para as inferências indutivas – o que não significa, em absoluto, que o Teorema de Bayes tenha perdido sua utilidade para explicar o raciocínio científico. Howson (2003), por exemplo, afirma que o enfoque bayesiano deve ser considerado não tanto uma teoria positiva da confirmação, mas como um arcabouço para implementar qualquer teoria da confirmação favorecida pelo analista.

Há outro exemplo mencionado por Howson e Urbach e devido a Rosencrantz, em que apenas a condição de equivalência é mantida, mas não a condição de Nicod. Também aqui a explicação bayesiana permite compreender a contento o que se passa. Suponha-se a hipótese "todos os gafanhotos são encontráveis fora do Condado de Yorkshire" (Howson & Urbach, 2006, pg. 102). Essa hipótese é equivalente a "nenhum gafanhoto é encontrável dentro do Condado de Yorkshire". Imagine-se agora que um gafanhoto foi encontrado exatamente na fronteira do condado. Esse indício constitui uma instanciação da hipótese e a confirmaria conforme a condição de Nicod. O problema é, porém, que como não existe controle de barreiras para gafanhotos, é possível dizer que se encontramos um gafanhoto nos exatos limites do Condado, é porque, na verdade, há um processo de migração para dentro dele. Embora não haja ainda nenhum gafanhoto dentro do Condado, a observação de um gafanhoto na linha divisória, em lugar de confirmar a hipótese original, na verdade a infirma. Contrariamente à condição de Nicod, a observação do gafanhoto, apesar de confirmar que nenhum gafanhoto é encontrável dentro do Condado, infirma a hipótese de que todos os gafanhotos são encontráveis fora do Condado, uma vez que há um processo de migração em curso. Em termos bayesianos, o que ocorre, nesse exemplo, é que a probabilidade de certo evento *e* é reduzida com relação à hipótese original *h*, isto é, $P(e/h) < P(e)$. Nessa situação, a hipótese *h* ("todos os gafanhotos são encontráveis fora do Condado de Yorkshire") é infirmada, isto é $P(h/e) < P(h)$.[13]

Em suma, apesar da crítica pontual de Strevens, a análise bayesiana consegue dar conta não apenas do paradoxo dos corvos, mas de várias situações similares com relação às quais outros enfoques enfrentam grandes dificuldades. Ao fazê-lo, chega a conclusões que se mostram mais condizentes com o senso comum. Isso demonstraria, insistem Howson e Urbach, sua vantagem com relação a enfoques alternativos.

[13] Há um terceiro exemplo, que me foi mencionado oralmente pelo Professor António Zilhão, da Universidade de Lisboa, que confirma o fato de que a condição de Nicod não é universalmente válida. Trata-se do exemplo de três amigos que vão a um restaurante. Tome-se, como hipótese, a afirmativa de que "ao sair de restaurante, cada um saiu com um chapéu que não é o seu". Nessas condições, um observador vê que A saiu com o chapéu de B, o que confirma a hipótese. O mesmo observador também vê que B saiu com o chapéu de A, o que resulta em nova instância de confirmação da hipótese. O problema é que a conjunção das duas sentenças isto é, $A \wedge B$ ("A e B"), em lugar de confirmar a hipótese, na verdade a infirma, pois logicamente decorre de tal conjunção que C saiu com o próprio chapéu. Trata-se de uma situação em que duas assertivas separadamente confirmam a hipótese original, mas sua conjunção a infirma, o que contraria a condição de Nicod.

O PROBLEMA DE DUHEM

A partir desta seção, intensifica-se, na exposição que desenvolvem Howson e Urbach, o diálogo com o falsificacionismo popperiano. De fato, a rejeição de Popper à possibilidade de o conhecimento científico repousar sobre bases indutivas fez que este se tornasse um dos mais acirrados críticos de interpretações subjetivistas da probabilidade e, por extensão, do bayesianismo. Nas seções que se seguem, Howson e Urbach persistem em seu objetivo de demonstrar que o bayesianismo é uma alternativa superior ao falsificacionismo na análise do raciocínio científico. Para isso, eles de debruçam sobre uma série de aspectos importantes da atividade científica, entre os quais o problema de Duhem.

Preliminarmente, porém, Howson e Urbach fazem uma crítica de cunho mais geral ao falsificacionismo de Popper. Tal crítica põe em questão, justamente, o traço mais significativo da visão popperiana da ciência, segundo o qual uma teoria só pode ser considerada científica se ela for falsificável, isto é, se puder ser refutada por observações empíricas efetivas ou, ao menos, concebíveis. O problema é que, ciente do fato de que não existem observações empíricas definitivamente conclusivas, Popper se viu forçado a reconhecer, para fins de falsificação, que as observações constituem, no mais das vezes, o resultado de convenção ou acordo de parte da comunidade científica (Howson & Urbach, 2006, p. 105).[14] Além disso, grande parte das teorias comumente consideradas como grandes realizações da ciência não satisfazem o critério estrito de Popper, sendo dificilmente falsificáveis nos moldes aludidos. Tais teorias fazem predições apenas com o auxílio de teorias ou hipóteses secundárias.

A segunda crítica feita por Howson e Urbach a Popper é que, ao se defrontar com uma observação falsificadora, é sempre possível ao cientista afirmar que o problema não está na teoria principal, mas em uma ou outra teoria ou hipótese auxiliar. Essa crítica a Popper pode ser sintetizada no chamado problema de Duhem: uma vez que teorias distintas foram utilizadas para fazer uma previsão que se mostrou falsa, qual dentre elas deve ser considerada como efetivamente falsificada?

A dificuldade de explicar de forma satisfatória o problema de Duhem não ficou restrita apenas a Popper. Tanto Kuhn (1970) quanto Lakatos (1970) tentaram explicar o chamado problema de Duhem. No caso de Kuhn, a teoria principal que o pesquisador

[14] "From a logical point of view, the testing of a theory depends upon basic statements whose acceptance or rejection, in its turn, depends upon our decisions. Thus it is decisions which settle the fate of theories" (Popper, K, 1959, p. 108).

insiste em conservar, a despeito das instâncias de falsificação, corresponderia ao chamado “paradigma científico”[15]. Para Lakatos, tratar-se-ia do chamado “núcleo duro” do programa de pesquisa científica, ao passo que as teorias e hipóteses auxiliares corresponderiam ao chamado “cinturão protetor”. Para Lakatos, na medida em que o “núcleo duro” se mostra ainda capaz de fazer novas previsões, o cientista está justificado em preservá-lo diante de instâncias de refutação. Nesse caso, o programa ainda poderia ser considerado progressivo. Caso contrário, estaria em fase de degeneração, o que pode levar, em última instância, a seu abandono.[16]

Lakatos não foi capaz de indicar em que momento um programa pode ser considerado em degeneração ou quando pode, apesar das evidências em contrário, ainda ser considerado progressivo (afinal, quantas novas previsões são necessárias para que o cientista esteja justificado em sua insistência de proteger o núcleo duro de seu programa de pesquisa?). Tampouco fica claro por que determinadas teorias recebem o status de “núcleo duro”, enquanto outras são consideradas parte do “cinturão protetor”. Fica-se, assim, com a impressão de que o cientista tem total liberdade ao fazer tal classificação, inexistindo parâmetros racionais que o guiem nessa tarefa.

O enfoque bayesiano consegue dar conta do problema de Duhem, o que seria uma clara mostra de sua superioridade aos enfoques de Popper, Kuhn e Lakatos. No caso, a aplicação do Teorema de Bayes evidencia como as respectivas probabilidades da teoria principal e das teorias auxiliares são modificadas de maneira distinta diante de uma instância de falsificação. O bayesianismo explicita o impacto assimétrico exercido por um indício com relação à hipótese principal e às hipóteses secundárias. Howson faz uso de um exemplo concreto para mostrar que a probabilidade posterior da teoria principal é afetada apenas de forma marginal. A probabilidade do chamado “cinturão protetor”, por sua vez, é afetada de forma substancial.

---

[15] Masterman (1979) constatou a ambiguidade apresentada pela palavra “paradigma”: o termo fora utilizado por Kuhn de vinte e duas maneiras diferentes. Reconhecendo as confusões induzidas pela apresentação original, Kuhn esclarece seu significado no Posfácio à *Estrutura das Revoluções Científicas*, edição de 1969. O termo paradigma tem um sentido geral e um sentido restrito. Em sentido geral, pode-se assimilar a noção de paradigma à noção de sistema de ideias. Em sentido restrito, porém, deve-se identificar a noção de paradigma com a expressão matriz disciplinar. Essa é composta por elementos ordenados de várias espécies, que são possuídos de forma coletiva pelos praticantes de determinada disciplina. Esses elementos são: generalizações simbólicas, modelos particulares, valores compartilhados e soluções experimentais específicas. Ver, a propósito, Masterman, M. *A natureza de um paradigma*. In: LAKATOS, I.; MUSGRAVE, A. *A crítica e o desenvolvimento do conhecimento*. São Paulo, Cultrix, 1979.

[16] Lakatos, Imre. Falsification and the Methodology of Scientific Research Programmes. In: *Criticism and the Growth of Knowledge.* Cambridge: Cambridge University Press, 1970.

No exemplo histórico de Howson, a teoria principal *t* constitui a chamada hipótese de Prout (que afinal acabou sendo descartada anos mais tarde), segundo a qual "o peso atômico de um elemento qualquer é, necessariamente, um número inteiro e múltiplo do peso atômico do hidrogênio". As hipóteses ou teorias auxiliares podem ser assimiladas à premissa *a* de que as técnicas de medição são precisas (dentro de certos limites) e as amostras empregadas estão livres de quaisquer impurezas. A instância falsificadora *e* é dada pelo peso atômico da clorina, que em 1815 foi medido em 35.83, valor incompatível com *t* e com *a*.

A hipótese de Prout era aceita como verdadeira pela maior parte dos químicos por volta de 1860. Por essa razão, Howson atribui a P(t) o valor de 0,9. Já a confiança nas técnicas de medição e na pureza das amostras era menor, razão pela qual se atribui a P(a) o valor de 0,6. Howon insiste em que tais probabilidades prévias, ainda que arbitrárias, não são inverossímeis do ponto de vista do contexto histórico de seu exemplo. Ao mesmo tempo, os valores atribuídos não são importantes, dentro de certos limites, para o objetivo buscado, que é o de lançar luz sobre a evolução histórica da hipótese de Prout, bem como evidenciar que teoria principal e hipóteses secundárias são afetadas de modo distinto por instâncias de falsificação.

Howson assume, baseado em Dorling (1979), que *t* e *a* são completamente independentes (a assunção de independência torna os cálculos mais fáceis, mas não é imprescindível para o argumento). Tal premissa foi criticada por Bovens e Hartmann (2003), mas Howson insiste em que, à luz do caso concreto em análise, a assunção de independência está longe de ser descabida. As probabilidades posteriores de *t* e *a* dependem de P(e), P (e/t) e P(e/a). Temos, então, que

$P(e) = P(e/t)P(t) + P(e/{\sim}t)P({\sim}t)$

$P(e/t) = P(e/t\&a)P(a/t) + P(e/t\&{\sim}a)P({\sim}a/t)$

O fato de *t* e *a* serem independentes faz que P(a/t)=P(a) e, consequentemente, P(~a/t)=P(~a). Uma vez que a expressão *t&a* é refutada pelo indício *e*, temos que P(e/t&a) = 0. A partir daí, é possível deduzir que:

$P(e/t) = P(e/t\&{\sim}a)P({\sim}a)$ e, de forma análoga

$P(e/a) = P(e/{\sim}t\&a)P({\sim}t)$

$P(e/{\sim}t) = P(e/{\sim}t\&a)P(a) + P(e/{\sim}t\&{\sim}a)P({\sim}a)$

Para prosseguir em sua demonstração, é necessário atribuir valores a P(e/~t&a), P(e/~t&~a) e P(e/t&~a). Para isso, Dorling recorreu ao fato de que muitos químicos do século XIX acreditavam que a hipótese de Prout não era verdadeira, mas que as hipóteses auxiliares sim o eram. Para esses químicos, ao contrário do que acreditava Prout, a distribuição dos pesos atômicos se comportava conforme uma variável aleatória.

A estatística nos ensina que, em se tratando de variáveis aleatórias, em que a distribuição de probabilidades é uniforme, dois valores para determinado peso atômico separados por um intervalo de 0,01 terão a mesma probabilidade (isso é válido independentemente da amplitude escolhida para o intervalo). Se tomarmos *e* como a constatação de que o peso atômico de uma substância qualquer está, por exemplo, entre 35,825 e 35,835, bem como que a hipótese *t* é falsa, mas as hipóteses auxiliares agregadas como *a* são verdadeiras, então teremos que P(e/~t&a) = 0,01. No caso de se considerar tanto *t* quanto *a* falsas, não haverá qualquer influência com relação à distribuição da variável aleatória "peso atômico" e, assim, é possível assumir que P(e/~t&~a) também é 0,01. Por fim, Dorling (e Howson) atribuem a P(e/t&~a) o valor ligeiramente maior de 0,02. Isso ocorre porque, se a hipótese de Prout for verdadeira e as hipóteses auxiliares falsas, ao medir determinado número atômico, haverá a tendência, como quer a hipótese, de os valores se concentrarem mais próximos dos números inteiros, que constituem múltiplos do peso do hidrogênio. Assim, a distribuição dos valores não será uniforme, mas estará mais concentrada ao redor dos valores inteiros, o que justifica a atribuição de probabilidade maior.

Com base nesses valores, é possível passar para o cálculo das probabilidades necessárias à aplicação do modelo bayesiano e mostrar que a hipótese principal e as hipóteses secundárias são afetadas de formas distintas por uma mesma instância refutadora *e*. Assim, temos que

$P(e/{\sim}t) = P(e/{\sim}t\&a)P(a) + P(e/{\sim}t\&{\sim}a)P({\sim}a) = 0{,}01 \times 0{,}6 + 0{,}01 \times 0{,}4 = 0{,}01$

$P(e/t) = P(e/t\&{\sim}a)P({\sim}a) = 0{,}02 \times 0{,}4 = 0{,}008$

$P(e/a) = P(e/{\sim}t\&a)P({\sim}t) = 0{,}01 \times 0{,}1 = 0{,}001$

$P(e) = P(e/t)P(t) + P(e/{\sim}t)P({\sim}t) = 0{,}008 \times 0{,}9 + 0{,}01 \times 0{,}1 = 0{,}0082$

E, por fim, é possível calcular as probabilidades posteriores finais, da forma seguinte:

$$P(t/e) = \frac{P(e/t)}{P(e)} \times P(t) = \frac{0{,}008}{0{,}0082} \times 0{,}9 = 0{,}878$$

$$P(a/e) = \frac{P(e/a)}{P(e)} \times P(a) = \frac{0{,}001}{0{,}0082} = 0{,}073$$

Dos cálculos aduzidos, fica claro que determinada instância refutadora da hipótese de Prout afeta de forma distinta aquela hipótese e o conjunto de hipóteses ou teorias auxiliares. A probabilidade da hipótese principal, que originalmente tinha o valor de 0,9, passa a ter o valor de 0,878. Já no caso das hipóteses auxiliares, a probabilidade posterior é 0,073 (a probabilidade prévia era 0,6). Face ao indício e, a probabilidade da hipótese principal é um pouco reduzida, mas permanece em patamar que não afeta sua credibilidade. As hipóteses auxiliares, por outro lado, têm sua probabilidade diminuída a um nível em que deixam de ser sustentáveis. Diante de uma instância de refutação, a hipótese de Prout continua a ser mais provavelmente verdadeira do que falsa; as hipóteses auxiliares têm muito maior chance de ser falsas do que verdadeiras.

O acúmulo de sucessivas instâncias de refutação terá, entretanto, o efeito de erodir a probabilidade de uma hipótese e fazer com que seja abandonada por completo. Nessas condições, estaríamos diante da situação que Lakatos qualificou como um programa de pesquisa em degeneração e que, para Kuhn, corresponderia ao de crise de um paradigma. De fato, foi o que aconteceu com a hipótese de Prout, que estava completamente desacreditada por volta de 1860, em grande parte em função do acúmulo de medidas cada vez mais precisas de pesos atômicos, cujos valores se mostraram com ela incompatíveis (Howson & Urbach, 2006, p. 113).

A partir desse exemplo, a conclusão extraída por Howson e Urbach é que o Teorema de Bayes oferece um arcabouço adequado à solução do problema de Duhem, o que não acontece com metodologias alternativas que rejeitam o uso do cálculo de probabilidades.

O PROBLEMA DA EVIDÊNCIA

O impacto assimétrico entre os efeitos de uma instância de infirmação e outra de confirmação sobre a teoria foram enfatizados por Lakatos (1970, p. 137), para quem "o cientista é encorajado por uma resposta positiva da natureza, mas não se deixa desencorajar em face de uma resposta negativa". Embora a observação de Lakatos seja verdadeira, ela padece de qualquer justificativa na metodologia daquele autor. Um dos pontos positivos do bayesianismo é, justamente, explicar por que e em que circunstâncias esse impacto assimétrico ocorre.

Outra vantagem da análise bayesiana é deixar claro que se determinado indício é dedutível de uma hipótese *t* conjugada com hipóteses e teorias altamente questionáveis e expressas como a, então o grau de confirmação sobre a teoria original, quando o indício for verificado, será pequeno. É o caso das várias predições realizadas por Velikovsky, as quais nunca chegaram a impressionar cientistas sérios, mesmo no caso daquelas que foram efetivamente verificadas. Velikovsky previu, por exemplo, a existência de grandes quantidades de petróleo no planeta Vênus, tendo como base a hipótese de que vários fenômenos naturais do passado haviam sido causados por uma colisão entre a Terra e determinado cometa. A essa hipótese Velikovsky acrescentava hipóteses auxiliares pouco verossímeis, como a de que o referido cometa – que, segundo ele, viria a se transformar do planeta Vênus – era formado por hidrogênio e carbono e que tais elementos se converteram em petróleo por meio das descargas elétricas geradas pelo impacto.

O bayesianismo é igualmente capaz de explicar o fenômeno, que pode ocorrer na prática científica, dos dados serem “bons demais para serem verdadeiros”. Isso acontece quando os indícios obtidos empiricamente encaixam-se tão perfeitamente na hipótese que se quer provar, que o cientista chega a duvidar de sua veracidade. Em se tratando da hipótese de Prout, isso aconteceria no caso de relatos sucessivos de medidas de números atômicos como números inteiros, não obstante a perfeita consciência dos cientistas da época de que as técnicas de medição estavam longe de ser precisas e se sujeitavam a erros experimentais. A existência de hipótese alternativa suficientemente plausível, que explique a obtenção dos indícios melhor do que a veracidade da hipótese testada, esclarece a suspeita quanto aos indícios “bons demais para ser verdadeiros”.

Imagine-se, por exemplo, no caso da hipótese de Prout, que este tivesse manipulado os instrumentos de medição, de modo a obter números inteiros, situação que expressamos como $\sim a'$ (a' é a hipótese auxiliar de que as técnicas de medição não são absolutamente precisas e padecem de erros experimentais). Nesse caso, como os instrumentos foram manipulados, o efeito de confirmação de uma medição *e* independe de a hipótese de Prout (*t*) ser verdadeira ou falsa. Temos, assim, que

$$P(e/t\&\sim a') = P(e/\sim t\&\sim a')$$

Howson (2006, pp. 116-117) deduz, a partir dessas considerações, que, no caso em que estamos analisando,

$$P(t/e) = \frac{P(e/t)}{P(e)} \times P(t) = \frac{P(e/t\&{\sim}a')P({\sim}a')P(t)}{P\,e/(t\&{\sim}a')\,P({\sim}a')} = P(t)$$

Dessa forma, *e* não confirma *t* de forma significativa, no caso de se encaixar com perfeição na hipótese que está sendo testada. Ao ser "bom demais para ser verdadeiro", o indício *e* acaba por não ter efeito sobre a crença na veracidade de t. Tal conclusão depende, obviamente, da existência de hipótese alternativa que explique o indício recolhido de forma mais adequada. Na situação que estamos analisando, caso o método de obtenção das medições tivesse sido considerado preciso e todas as precauções tivessem sido tomadas contra manipulações de qualquer espécie, o poder de confirmação do indício *e* não mais estaria sob suspeita.

AS HIPÓTESES *AD HOC*

Nas seções precedentes, foi examinada a situação em que uma hipótese principal, combinada com outras, prediz a ocorrência de determinado indício. Uma vez que tal predição se mostra falsa, viu-se ser possível manter intacto o status epistêmico da hipótese principal ao creditar a falha a uma ou várias hipóteses auxiliares, que passam a ser desacreditadas, seja em parte, seja por completo. Ao lidar com essa questão, Hempel (1966, p. 29) propôs que se considerasse inferior, do ponto de vista epistêmico, uma hipótese ou teoria apresentada com esse fim, isto é, preservar uma hipótese que se encontra ameaçada frente a indícios contrários.

Como o bayesianismo se coloca frente às chamadas hipóteses *ad hoc*? Com base no bayesianismo, é possível demonstrar que certas hipóteses são *ad hoc* e, portanto, inadmissíveis, por maio da aplicação do Teorema de Bayes. Isso acontece porque se determinada hipótese for *ad hoc*, sua propositura não resultará no aumento da probabilidade posterior da hipótese original. Um indício *e* que passa a ser previsto com base na conjunção da hipótese *h* com a hipótese "ad hoc" *a´* terá um impacto de confirmação desprezível sobre a hipótese original.

A única forma de avaliar hipóteses e concluir se elas são ou não *ad hoc* é, para Howson e Urbach, por meio do Teorema de Bayes. Todas as tentativas de alcançar esse objetivo por outros enfoques, como os propostos por Popper, Lakatos e Hempel, se mostram equivocadas. Em particular, ao exigirem que determinada hipótese, para não ser *ad hoc*, produza indícios independentes, Popper, Lakatos e Hempel não apenas não solucionam o problema da identificação das hipóteses *ad hoc*, como dão margem para o surgimento de situações claramente anti-intuitivas. Por meio da aplicação do Teorema de Bayes, Howson e Urbach mostrarão que a utilização do conceito de independência para identificar hipóteses "ad hoc", seja como independência lógica, seja como independência probabilística, conduz a resultados indesejáveis.[17]

Em determinadas situações, a propositura de hipóteses *ad hoc* se mostra claramente inadequada. Howson menciona o exemplo de ambientalistas que atribuíam o desempenho individual em testes de QI exclusivamente a fatores ambientais. Em face de um grupo de esquimós cujos resultados nesse gênero de testes se mostraram particularmente exitosos, mas que levavam uma vida absolutamente desregrada e irregular (falta de trabalho fixo, abuso da bebida, etc.), o biólogo Peter Medawar alegou que "a vida nos iglus fornece a combinação adequada de proteção, segurança e contato mútuo que levam a bom desempenho nos testes de inteligência" (Howson e Urbach, 2006, p. 120).

A situação acima difere da propositura da hipótese de existência de um novo planeta – que mais tarde descobriu-se ser, de fato, Netuno – para explicar as anomalias na órbita de Urano. Nessa segunda situação, a hipótese proposta, que não é *ad hoc*, permitiu salvar a Teoria da Gravitação de Newton de instância de refutação oriunda das falhas quanto às predições relativas ao movimento e órbita de Urano.

Qual seria a diferença entre os dois exemplos acima? Uma maneira possível de compreender tal diferença é por meio da noção de indícios independentes. Quando determinada hipótese *t* e as respectivas hipóteses auxiliares *a* são refutadas por um indício *e*, não é suficiente substituir as hipóteses auxiliares por outras – por exemplo, *a´*–, de modo que aquele indício se torne consequência da proposição *t&a´*. É preferível que a proposição *t&a´* não apenas dê conta do indício que refutava *t&a*, mas que também seja capaz de levar a predições independentes e verificáveis. Tal critério foi proposto por Bacon e, mais recentemente, por Popper (1963, p. 241) e Lakatos (1970, p.

[17] Diz-se que as proposições A e B são logicamente independentes se uma não decorre logicamente da outra. A noção de independência probabilística, por sua vez, refere-se a outro fenômeno. Na teoria das probabilidades, dois eventos são probabilisticamente independentes quando a probabilidade de um deles ocorrer não é influenciada pela probabilidade de o outro ocorrer, isto é $P(A/B) = P(A)$ *e* $P(B/A) = P(B)$.

175), os quais coincidem em que nova teoria deve ser A de forma independente. Howson se oporá a esse critério e contraporá a ele o uso do Teorema de Bayes.

O problema apontado por Howson é que o requisito de que a nova hipótese ou teoria deva gerar novas predições ou, como prefere Popper, seja testável independentemente, nem sempre se sustenta diante de exemplos da prática científica. Em muitas ocasiões, a hipótese revisada é perfeitamente plausível sem que ela produza ou existam novos indícios de confirmação. Um exemplo citado por Howson é o da retirada de bolas de uma urna em que, acredita-se, só há bolas brancas. Suponha-se que as bolas são retiradas ao acaso e que, após cada retirada, a cor é anotada e a bola reintroduzida na urna. Após 10.000 retiradas, obtiveram-se 4.950 bolas vermelhas e 5.050 bolas brancas. É natural inferir, diante desse conjunto de indícios, que a urna contém bolas brancas e vermelhas em quantidades aproximadamente iguais. Tal inferência seria uma hipótese revisada e plenamente justificada pelos dados empíricos obtidos, muito embora não exista qualquer evidência independente para ela. A hipótese original ("a urna contém apenas bolas brancas") foi revisada ("a urna contém bolas brancas e vermelhas em quantidades aproximadamente iguais") tendo, como justificativa, o conjunto de indícios já produzidos e nenhum indício futuro.

O segundo exemplo aduzido deixa mais claro o motivo pelo qual a exigência das predições independentes nem sempre é critério adequado para identificar uma hipótese como "ad hoc". De acordo com esse exemplo, supõe-se que duas características de certa planta são herdadas por meio de um par de genes independentes e localizados em cromossomos distintos. Uma série de experimentos resulta, porém, em plantas que apresentam simultaneamente os dois tipos fenotípicos, o que leva à rejeição da hipótese de independência, em favor da hipótese de que ambos os genes estão ligados no mesmo cromossomo. Assim como na situação anterior, a hipótese anterior é descartada apenas com base nas informações empíricas já produzidas e sem que haja qualquer necessidade de novos e independentes indícios.

Howson e Urbach (2006, p. 123) também indicam uma terceira objeção à exigência de predições independentes como critério para classificar hipóteses como "ad hoc". Considere-se um cientista que quer avaliar o status epistêmico das hipóteses *t&a*. Esse cientista realiza um experimento na esperança de obter o indício *e* mas, no lugar dele, obtém *e´*, que é incompatível com as hipóteses originais. Esse cientista modifica as hipóteses originais para *t&a´*, que passam a explicar *e´*, mas não acarretam qualquer nova predição. Tome-se, agora, o caso de um segundo cientista, também interessado em testar *t&a*, com a diferença de que seu experimento somente pode produzir *e* ou *~e*. Ao obter *~e*, o segundo cientista reformula suas hipóteses originais para *t&a´*. Em consequência dessa reformulação, efetua a predição do indício *e´*, que se mostra bem-sucedida.

O problema do exemplo acima é que chegamos a conclusões diametralmente opostas com base nos mesmos indícios. No primeiro caso, *t&a´* padece dos males de uma hipótese "ad hoc"; no segundo caso, é uma hipótese bem sucedida, uma vez que prediz, com êxito, o indício *e´*. A explicação para tal anomalia está na ordem entre teoria e indícios para cada um dos cientistas. No primeiro caso, *t&a´* foi concebida após a identificação de *e´*; no segundo caso, a ordem foi inversa. Levar em consideração fatores como a ordem em que os cientistas concebem uma hipótese com relação a determinada instância de refutação ou confirmação introduz novas e inusitadas complicações na tentativa de analisar o raciocínio científico, inclusive quanto ao estado mental do cientista, o que é difícil de compatibilizar com uma visão objetiva da ciência. A insistência em que uma teoria produza indícios independentes para não ser "ad hoc", por si só, não resolve a questão e pode levar a situações estranhas como a que foi descrita. O enfoque bayesiano, por outro lado, é capaz de explicar porque as pessoas reagem com incredulidade quando tais hipóteses são apresentadas, sem que isso leve às dificuldades assinaladas.

Toda a crítica tradicional às hipóteses "ad hoc" baseia-se no argumento de que elas serão dificilmente aceitáveis, a menos que apresentem algum indício novo e independente. Howson e Urbach (2006, p. 126) argumentam, por outro lado, que a única maneira de se ter uma ideia precisa acerca de como esses indícios podem confirmar a hipótese reformulada é através de critérios bayesianos e não através da noção vaga de independência.

Suponha que determinada hipótese foi refutada pelo indício *e´*. A hipótese reformulada *h* passa a explicar e predizer *e´*, bem como o novo indício *e´´*. Não há dúvida que *e´´* é uma instância de corroboração de *h*. Ao utilizar o Teorema de Bayes, assumindo-se que *e´´* é um indício suficientemente improvável, tem-se que P(e´´/h&e´) > P(e´´/e´). A aplicação do Teorema de Bayes deixa claro, assim, que, se o indício for suficientemente improvável, ele terá o condão de confirmar a hipótese em exame. Isso é feito por meio da aplicação do Teorema e do cálculo de probabilidades e, não, por meio da noção vaga de independência do indício.

O efeito confirmatório de *e´´* sobre *h* pode se dar mesmo no caso de *e´* e *e´´* não serem probabilisticamente independentes (em probabilidade, dois eventos A e B são independentes se a probabilidade de um deles ocorrer não depender do fato de o outro ocorrer), o que mais uma vez evidencia o equívoco de se recorrer a essa noção. Por meio do Teorema de Bayes, seria possível mostrar que haverá efeito confirmatório de *e´´* sobre *h*, mesmo no caso de *e´´* e *e´* não serem independentes. O ponto é que esse efeito confirmatório seria pequeno. O simples recurso à noção de independência, como exigido por Popper e Lakatos, não conseguiria explicar porque isso acontece.

Uma tentativa de salvar o critério de Popper e Lakatos é por meio da afirmativa de que, ao exigirem indícios independentes para a hipótese não ser "ad hoc", eles na verdade estão se referindo apenas à noção de independência lógica e não independência probabilística. Seria possível, assim, pensar a exigência de um indício novo e independente do ponto de vista apenas lógico, sem implicar independência probabilística. Ocorre, porém, que isso tampouco resolveria o problema, pois embora *e´´* seja distinto e logicamente independente de *e´*, não decorre daí que aquele indício produzirá qualquer impacto do ponto de vista da confirmação de *h*. Imagine-se, por exemplo, que a diferença entre *e´* e *e´´* é apenas um elemento trivial, como local ou momento de ocorrência. Estar-se-ia diante de dois eventos que resultam de predições de *h*, muito embora o segundo indício não tenha qualquer impacto de confirmação sobre a hipótese. Isso ocorre porque, para que um novo indício tenha um efeito de confirmação, é intuitivamente necessário que seja significativamente distinto dos indícios anteriores e não distinto apenas em um sentido trivial.

Também nesse exemplo a análise bayesiana nos permite compreender o que está em jogo. Se os indícios *e´* e *e´´* forem similares, $P(e´´/e´) \cong 1$, o que implica que o impacto confirmatório de *e´´* será ínfimo. No caso de serem distintos, então $P(e´´/e´) < 1$, o que implica que *e´´* exerce um impacto de confirmação sobre *h*.

O bayesianismo esclarece as circunstâncias e as condições necessárias para que um novo indício decorrente de uma hipótese reformulada tenha um efeito de confirmação, o que permite conservar aquela hipótese em bases racionais e afirmar que ela não é "ad hoc". A exigência de que o novo indício deva ser independente probabilisticamente tem consequências duvidosas, já que indícios não independentes nesse sentido podem ter impacto de confirmação, embora pequeno, conforme a análise bayesiana deixa claro. Ao mesmo tempo, a exigência de independência não pode simplesmente ser equiparada à independência lógica, pois nesse caso, por outro lado, não haveria necessariamente confirmação. A conclusão de Howson é que a tentativa de solucionar o problema de Duhem com base em que se devam evitar hipóteses “ad hoc”, a menos que estas produzam indícios novos e independentes, é repleta de falhas. Apenas a análise bayesiana é capaz de esclarecer os pontos obscuros de análise. Através da análise bayesiana, é possível verificar o impacto de um indício novo com relação à hipótese original e às hipóteses reformuladas. Se esse impacto for confirmatório e seu valor não for desprezível, então está claro que a hipótese reformulada não é “ad hoc”. A análise bayesiana permite chegar a esse resultado sem cair nas armadilhas a que está sujeito o critério proposto por Popper, Lakatos e Hempel.

CONCEPÇÃO DE EXPERIMENTOS

Em *Scientific Reasoning – the Bayesian approach*, Howson e Urbach apresentam uma exposição sistemática do enfoque bayesiano para interpretar o raciocínio científico. Ao longo deste capítulo, procurei expor as ideias centrais, da forma como foram elaboradas por aqueles dois autores, quanto às vantagens do enfoque que propõem em comparação a outros modelos, como os de Popper e Lakatos. Ao final do capítulo 4 do livro, que é o capítulo principal no que se refere à exposição do bayesianismo, Howson e Urbach analisam uma questão particularmente importante. Trata-se da crítica, originalmente endereçada a Popper e posteriormente estendida ao bayesianismo, de que nenhum dos dois enfoques explica adequadamente as razões por que os cientistas se preocupam tanto em realizar novos experimentos e reunir mais indícios.

Elaborada por Maher (1990), a crítica tem a seguinte formulação: uma vez que nenhum indício pode corroborar ou refutar definitivamente uma hipótese, Popper não é capaz de justificar as razões dos cientistas realizarem novos experimentos para obter mais informações sobre o mundo. O argumento, estendido ao bayesianismo por Miller (1991), toma a forma seguinte: tome-se *e* pelo conjunto total de indícios à disposição do pesquisador e P(h/e) como a probabilidade posterior da hipótese *h*. Quais são os incentivos para que esse pesquisador realize novos experimentos com vistas a modificar essa probabilidade, uma vez que a hipótese jamais será confirmada ou infirmada definitivamente? Sem dúvida, ele pode realizar novas experiências e modificar seu conjunto de indícios de *e* para *e´*, mas é difícil compreender porque o faria, uma vez que P(h/e´) não confirma ou infirma *h* melhor do que P(h/e)[18].

A resposta a essa observação é que o propósito de uma investigação científica, nos moldes defendidos pelo bayesianismo, não é avaliar melhor probabilidades indutivas, mas diminuir o grau de incerteza sobre determinado aspecto do mundo. A crítica de Miller se fundaria, assim, em uma compreensão equivocada dos propósitos do enfoque bayesiano e seus fins precípuos.

Tome-se o exemplo do cientista que está interessado em conhecer determinada variável. A princípio, é possível que esse cientista não saiba ao certo como essa variável se comporta, uma vez que seus valores possíveis são excessivamente difusos. É possível, porém, diminuir esse grau de incerteza por meio de um experimento que modifique a distribuição de probabilidade de tal variável. Com o auxílio do Teorema de Bayes, é possível obter nova distribuição, mais concentrada em determinada região, o que permite ao cientista conhecer melhor a variável.

É verdade que o Teorema de Bayes não indica a maneira de construir um experimento que se mostre útil para diminuir o grau de incerteza da variável que se quer conhecer. A relação relativamente indiferente do bayesianismo com as técnicas experimentais será, a propósito, um dos pontos que Deborah Mayo ressaltará ao defender as vantagens comparativas da *error statistics*. Um experimento bem concebido pode se mostrar decepcionante do ponto de vista das informações que produz e, inversamente, um experimento mal concebido pode produzir indícios surpreendentemente úteis.

Howson e Urbach admitem a pertinência da crítica acima, mas para eles ela não se mostra um grande problema, já que exige do bayesianismo algo que está fora de seus propósitos. Escrevem Howson e Urbach a esse respeito: “Ao decidir se é conveniente realizar determinado experimento, ao menos três fatores devem ser levados em conta: o custo do experimento; a moralidade de realizá-lo; e o valor, tanto de um ponto de vista teórico quanto prático, das hipóteses em que se está interessado. O Teorema de Bayes,

[18] A questão da refutação em definitivo de uma teoria nem sempre é isenta de controvérsias na Filosofia da Ciência. De fato, é possível dizer que a Teoria de Newton foi definitivamente refutada pela Teoria Geral da Relatividade? Muitos se apressarão a responder afirmativamente a essa pergunta, embora isso não explique as razões da Teoria Newtoniana continuar a ser ensinada, ainda hoje, em universidades e cursos superiores não apenas como uma curiosidade histórica, mas como teoria que permite a resolução de um grande número de problemas reais. Talvez a refutação seja uma questão de contexto: no contexto de velocidades próximas à da luz, a Teoria de Newton está completamente refutada; no contexto terrestre, porém, ela permanece válida. Essa é uma questão para a qual não tenho resposta definitiva e que merece ser aprofundada, embora escape ao objetivo desta dissertação.

obviamente, não pode ajudar aqui” (Howson & Urbach, p. 127). Diferentemente de Popper, para o bayesianismo, o objetivo da ciência não é tentar falsificar hipóteses, mas diminuir o grau de incerteza sobre o mundo. Nessa empreitada, que tipo de experimentos será ou não empregados é uma questão desprovida de relevância.

INDETERMINAÇÃO E PROBABILIDADES PRÉVIAS

Ao longo das seções precedentes, muito se falou acerca do problema de Duhem, ou da indeterminação, e das formas como o bayesianismo propõe soluções que permitam lidar com ele. Há um elemento adicional relevante nessa questão: trata-se do exame das probabilidades prévias como forma de fazer frente a uma situação de indeterminação. Nesta seção, Howson e Urbach examinam a situação em que um mesmo indício e é explicado por hipóteses distintas e mutuamente excludentes. O bayesianismo oferece uma saída para o cientista que se encontra diante de tal dilema. Diante da necessidade de escolher uma dentre várias hipóteses que explicam um mesmo indício e, o cientista poderá recorrer ao valor da probabilidade prévia como critério de escolha, não obstante o valor da verossimilhança das várias hipóteses seja o mesmo. Para eles, também aqui o bayesianismo se consolida como alternativa superior ao falsificacionismo popperiano, que não indica nenhum método para discriminar entre hipóteses excludentes que expliquem da mesma forma o indício e.

É possível perguntar ao cientista bayesiano como ele lida, na atribuição de probabilidades prévias, com duas hipóteses que explicam de modo idêntico um determinado conjunto de indícios e. Jeffreys (1983) traz o exemplo de um conjunto infinito de hipóteses capazes de determinar corretamente a posição de um corpo em queda livre em determinado instante. Segundo o exemplo de Jeffreys, Galileu poderia ter apresentado a lei $S = a + u \times t + 1/2 \times gt^2 + f(t) \times f(t - t_1) \times f(t - t_2) \times ... \times f(t - t_n)$, que permite chegar à posição de um corpo em queda livre em determinado instante e que seria tão compatível com os resultados experimentais registrados quanto a equação efetivamente apresentada, isto é, $S = a + u \times t + 1/2 \times gt^2$. No caso, $t_1, t_2, ..., t_n$ constituem os instantes de queda que Galileu registrou em cada um de seus experimentos; a, u e g são constantes; e f é uma função qualquer que não seja infinita nos instantes $t_1, t_2, ..., t_n$. A equação de Jeffreys permite chegar a um conjunto infinito de alternativas à equação original de Galileu. Todas essas hipóteses alternativas são capazes de explicar os indícios registrados por Jeffreys, são mutuamente excludentes entre si e dão origem a predições distintas com relação a experimentos futuros.

O problema posto pelo exemplo de Jeffreys é a existência de um número infinito de hipóteses que se mostraram verdadeiras nas situações precedentes, mas que certamente não serão todas verdadeiras em situações futuras. Para os enfoques filosóficos que não recorrem ao Teorema de Bayes, inclusive o enfoque popperiano, tanto a equação de Galileu como as equações de Jeffreys devem ser consideradas alternativas igualmente válidas ao se considerarem os indícios disponíveis. Tal conclusão contradiz, todavia, a noção intuitiva de que os cientistas, ao cotejar hipóteses distintas que explicam um mesmo conjunto de indícios, naturalmente consideram algumas mais plausíveis de que outras.

De acordo com a análise bayesiana, o fato de duas ou mais hipóteses explicarem igualmente bem um conjunto de indícios implica, tão somente, que suas verossimilhanças posteriores são idênticas, isto é, $P(e/h_1) = P(e/h_2)$. Isso não significa, em absoluto, que as probabilidades prévias sejam iguais. Se voltarmos ao exemplo de Jeffreys, é perfeitamente plausível que o cientista confrontado com as várias equações distintas atribuiria uma probabilidade prévia mais alta à equação de Galileu e probabilidades prévias mais baixas às hipóteses concorrentes.

As probabilidades prévias exercem, assim, um papel importante para explicar o raciocínio científico, na medida em que permitem compreender as razões pelas quais os cientistas, muitas vezes, preferem uma hipótese que explique determinado indício de modo imperfeito, em que $P(e/h) < 1$, a outras que o expliquem melhor. Isso acontece porque a capacidade explanatória supostamente melhor das hipóteses alternativas é anulada por probabilidades prévias mais baixas, que permitem, inclusive, pensar tratar-se de hipóteses *ad hoc*.

O enfoque bayesiano não oferece, porém, um método geral para a atribuição de probabilidades prévias. A acusação de subjetivismo é, com efeito, uma das críticas mais fortes ao bayesianismo, como veremos adiante, não obstante as respostas possíveis ensaiadas por seus defensores, por exemplo, quando indicam o papel da simplicidade da hipótese na atribuição da probabilidade prévia. O edifício bayesiano repousa, em grande parte, sobre o fundamento de que a atribuição da probabilidade prévia constitui a expressão mesma da atitude epistêmica do cientista com relação a sua teoria científica. Apesar de toda sua articulação interna, permanece uma questão aberta, do ponto de vista do argumento bayesiano, se essa atribuição constitui um elemento descritivo do raciocínio científico ou um elemento prescritivo.

A LEI DE ADIÇÃO DE PROBABILIDADES

Ao longo das três edições de *Scientific Reasoning*, Howson e Urbach depuraram continuamente a apresentação do enfoque bayesiano para o raciocínio científico. Embora os aspectos principais tenham sido mantidos a cada reformulação, há diferenças cuja explicitação é necessária, inclusive como forma de melhor conhecer a evolução do bayesianismo à medida que seus argumentos foram submetidos ao exame crítico de filósofos e cientistas.

Um primeiro ponto digno de registro é que, na edição de 2006 de *Scientific Reasoning*, Howson e Urbach passaram a aceitar os conhecidos argumentos elaborados por De Finneti contra o chamado princípio ou lei da adição de probabilidades (*principle of countable additivity*). Embora não afete a apresentação do argumento bayesiano, a modificação de posição é relevante, uma vez que se trata de um princípio que, do ponto de vista matemático, é considerado tão fundamental para o Cálculo de Probabilidades quanto os seus demais axiomas, muito embora possa ser questionado filosoficamente.

O segundo ponto, que será elaborado na seção seguinte, diz respeito à condicionalização com base no Teorema de Bayes. Na edição de 1993 de *Scientific Reasoning*, Howson e Urbach adicionaram um capítulo que estava ausente na edição de 1989 – e que veio a ser retirado na edição de 2006 -, no qual examinam uma série de questões fundamentais do ponto de vista da condicionalização. Ao examiná-las, o propósito de Howson e Urbach era, conforme ficará claro a seguir, mostrar que a análise que propõem é válida mesmo no caso em que a atualização de probabilidades posteriores não observe a regra da condicionalização bayesiana. Com isso, procuraram evidenciar que o enfoque bayesiano está fundamentado na ficção de um cientista ideal, não podendo ser tomado como uma descrição da prática diária nos laboratórios. O fato de que cientistas não se comportem como o investigador ideal bayesiano não invalida a aplicação do enfoque à prática da investigação científica.

Considerem-se dois eventos mutuamente exclusivos A_1 e A_2 associados com um experimento aleatório. Seja $A = A_1 \cup A_2$ a união desses dois eventos. Suponha que o experimento em questão seja repetido um grande número de vezes e produza, assim, uma série de tentativas sob condições idênticas. Seja *n* o número total de tentativas, e sejam $n(A_1)$, $n(A_2)$ e $n(A)$ o número de tentativas dos eventos A_1, A_2 e A, respectivamente. Se, para determinada tentativa, obtemos A, então isso significa que se obteve A_1 ou A_2, mas nunca ambos, uma vez que se trata de eventos mutuamente exclusivos. É evidente que

$$n(A) = n(A_1) + n(A_2)$$

e também que

$$\frac{n(A)}{n} = \frac{n(A_1)}{n} + \frac{n(A_2)}{n}$$

Ocorre que, para n suficientemente grande, cada um dos termos acima, que expressa a frequência relativa para A, A_1, A_2, passa a coincidir com as probabilidades respectivas P(A), P(A_1) e P(A_2). Nessa situação, segue que

$$\mathrm{P}(A) = \mathrm{P}(A_1) + \mathrm{P}(A_2)$$

De forma mais geral, é possível obter a seguinte formulação, que corresponde precisamente à chamada lei, princípio ou axioma da adição de probabilidades:

$$P\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} P(A_k)$$

A fórmula acima pode ser entendida da seguinte forma: se A_1, A_2, A_3, ... são um conjunto de proposições mutuamente excludentes no domínio P e a proposição “uma das proposições A_k é verdadeira” está também incluída em P, então a probabilidade total $P(A_k)$ é igual à soma $P(A_1) + P(A_2) + \cdots + P(A_n)$. Pode-se creditar a Kolmogorov boa parte do *status* atribuído à lei da adição das probabilidades. De fato, o matemático russo incluiu a referida lei como um dos axiomas fundamentais do cálculo das probabilidades em uma famosa monografia de 1950. Com isso, Kolmogorov fez do cálculo de probabilidades uma extensão da teoria matemática da medida, que se tornou desde então o paradigma dominante para entender probabilidades.

Howson e Urbach ressaltam que considerações matemáticas explicam, em grande medida, a adoção do axioma da adição como fundamento do cálculo de probabilidades. Sem esse axioma, seria extremamente difícil, por exemplo, apresentar certas versões do teorema do limite no cálculo de probabilidades. Eles asseveram, porém, que não obstante tais considerações, a razão fundamental para se adotarem axiomas é que eles permitem formular argumentos válidos. Independentemente da interpretação dada a esses axiomas, suas consequências devem ser, necessariamente, verdadeiras. No caso específico do axioma da adição das probabilidades, haveria, segundo Howson e Urbach, boas razões para considerá-lo, em determinadas situações, falso.

Ao se medir probabilidades ou tendências como frequências relativas que tendem a um limite, não há nenhuma razão para considerar verdadeiro o axioma, já que tais frequências relativas, diferentemente das frequências em amostras limitadas, nem sempre estão a ele sujeitas. Especificamente, no caso de determinadas ocorrências que tendem a acontecer apenas um número limitado de vezes, o limite da frequência correspondente é zero no caso de o número de tentativas tender ao infinito.

De Finetti indicou (1972, p. 86) ser perfeitamente razoável, sob determinadas condições, atribuir probabilidade zero a cada um dos membros de uma soma infinita de partições exaustivas do universo total de possibilidades, embora isso vá de encontro ao axioma, uma vez que a probabilidade total do universo de possibilidades deve ser 1.

O argumento de Howson (2006, p. 28) é, assim, que em diversas ocasiões há boas razões para atribuir o valor uniforme zero a cada um dos resultados possíveis e mutuamente exclusivos de um universo de possibilidades. Tal possibilidade é, porém, vedada pelo axioma, o que faz que o analista seja forçado a adotar uma distribuição de probabilidades equivocada. Na verdade, ele sublinha que o axioma somente é claramente aplicável no caso de as probabilidades da partição do espaço de resultados possíveis formar uma sequência que converge para o valor 1, como, por exemplo, a progressão $\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots$

Que motivo teria levado Howson e Urbach a introduzir uma seção específica em *Scientific Reasoning* para rejeitar a lei da adição de probabilidades? Uma primeira explicação é que a rejeição do axioma constitui uma evolução importante do ponto de vista teórico que aproxima Howson e Urbach ainda mais de De Finnetti. De Finnetti foi um dos principais idealizadores da interpretação subjetivista das probabilidades. Seus escritos são a base sobre a qual repousa a exposição do bayesianismo levada a cabo por Howson e Urbach.

Há, porém, uma segunda razão para esta seção, que é mais relevante. Vários bayesianos, como Jaynes e Williamson, nunca estiveram completamente à vontade com a interpretação subjetivista das probabilidades, tendo buscado identificar critérios "objetivos" para a atribuição das probabilidades prévias.[19] Um desses critérios é, precisamente, o axioma da adição das probabilidades. Ao rejeitar esse axioma, Howson e Urbach estão dando um primeiro passo para negar a possibilidade de constranger a liberdade do investigador para atribuir probabilidades prévias. Com isso, eles reafirmam seu entendimento da noção de probabilidade como valor de crença e, mais especificamente, como disposição do pesquisador para apostar na veracidade da hipótese, como proposto por De Finnetti.

CONDICIONALIZAÇÃO BAYESIANA

O edifício bayesiano repousa na tese de que os graus de crença do pesquisador são consistentes, o que significa, para fins práticos, que eles obedecem os quatro axiomas fundamentais do cálculo de probabilidades:

(1) $P(a) \geq 0$ para todo a no domínio de P;

(2) $P(t) = 1$ se t é uma tautologia;

(3) $P(a \vee b) = P(a) + P(b)$ se a , b e $a \vee b$ estão no domínio de P, e a e b são mutuamente exclusivos; isto é, o valor de verdade V atribuído a um deles implica o valor de verdade F atribuído ao outro;

(4) $P(a/b) = P(a\&b)/P(b)$

Uma das consequências dos axiomas acima é, precisamente, o Teorema de Bayes, o qual regula a forma como os graus de crença são atualizados face à ocorrência de determinado indício *e*. Atualizar graus de crença significa abandonar determinada função de probabilidade P, com relação à qual o indício *e* apresenta uma probabilidade menor do que 1, visto que anterior à sua ocorrência, por outra função de probabilidade P´, em que a probabilidade de *e* passa a ser 1. Ao longo desse processo, proposições no domínio de P adquirirão novos valores em P´. Coloca-se, então, a questão de saber que valores serão esses, isto é, como regular a passagem de P para P´. A resposta não poderia ser outra: a passagem de P para P´ dá-se precisamente conforme o Teorema de Bayes, isto é

$$P(h/e) = \frac{P(e/h)}{P(e)} \times P(h)$$

[19] Para se ter uma ideia do que constitui o bayesianismo objetivista, ver, por exemplo, Jaynes, E.T. Bayesian Methods: General Background. In *Maximum-Entropy and Bayesian Methods in Applied Statistics*, by J. H. Justice (ed.). Cambridge: Cambridge Univ. Press, 1986.

Isso significa, como é de se esperar, que $P'(h) = P(h/e)$. O Princípio da Condicionalização Bayesiana significa, então, que quando determinado grau de crença em e assume o valor 1, uma vez que o indício se verificou, o valor de $P'(a)$ deve ser atualizado para $P(a/e)$, para todo a no domínio de P, que é a função de probabilidade imediatamente anterior à atualização em questão. Embora tenha sido considerado por bom tempo como óbvio, o Princípio da Condicionalização Bayesiana nunca chegou a ser objeto de justificação formal, possivelmente porque se acreditava não haver necessidade para tal.

Surge, nesse contexto, a questão de saber se determinada estratégia de atualização de probabilidades que não obedeça ao princípio de condicionalização bayesiana leva, ou não, a inconsistências. Não há dúvida de que não é possível criar um caderno de apostas holandês no caso em que tal atualização ser feita através do Princípio de Condicionalização Bayesiana. A questão é, porém, se uma tal anotação poderia ser estabelecida no caso de a atualização de probabilidades ser efetuada por outros meios, isto é, no caso em que $P'(a) \neq P\,(a/e)$ e, ainda, se isso levaria a inconsistências.

Tal questão, longe de ser apenas teórica, revela-se de grande importância para o enfoque bayesiano. O ponto, aqui, é que dificilmente o Princípio da Condicionalização Bayesiana é seguido na prática diária dos laboratórios, quando cientistas comparam hipóteses rivais. Se for mostrado que formas de atualização de probabilidades que não obedeçam ao princípio podem levar a inconsistências, então a aplicação do enfoque bayesiano como instrumento para analisar o raciocínio científico estaria ameaçada, por ser válida apenas no contexto idealizado de cientistas cuja prática seja pautada forma rígida pela aplicação do Princípio da Condicionalização.

Em um influente artigo de 1967, mencionado por Howson e Urbach e intitulado *Slightly More Realistic Personal Probability*, Ian Hacking argumenta que não parece ser possível elaborar um caderno de apostas holandês se a atualização de probabilidades se der de forma consistente, mesmo que isso não ocorra por intermédio da condicionalização bayesiana. Tal conclusão foi, todavia, questionada alguns anos depois por Paul Teller que, inspirado em David Lewis, mostrou ser possível conceber um caderno de apostas holandês no caso de uma atualização de probabilidades não se dar por condicionalização (Teller, 1973).

Howson e Urbach reconhecem que o caderno de apostas holandês construído por Teller é válida, mas asseveram que ela não implica que apenas a condicionalização bayesiana torna a atualização de probabilidades imune contra inconsistências. O exemplo de Teller é aplicável apenas no caso de um conjunto de graus de crenças que são consideradas simultaneamente racionais (caderno de apostas holandês diacrônico) e cuja atualização não se dá por condicionalização. Em um caso em que os graus de crença variam com o tempo e são atualizados de forma consistente por outros meios, o argumento de Teller não se aplica.

A conclusão de Howson e Urbach é que, embora existam situações em que a condicionalização – seja por meio do Teorema de Bayes, seja por outros meios – é necessária, ela não é a única maneira válida de atualizar probabilidades de modo a evitar inconsistências. De acordo com o enfoque para probabilidades esposado pelo bayesianismo, a probabilidade prévia $P(h)$ corresponde à estimativa de valor que o investigador está disposto a apostar na veracidade de h, relativamente ao estado K de seu conhecimento atual. Nessas condições, a probabilidade condicional $P(h/e)$ corresponde à estimativa do valor que o investigador está disposto a apostar na verdade de h com relação ao estado de conhecimento $K \cup \{e\}$. Com base na suposição de que, ao tomar conhecimento da ocorrência de e, isto é, ao adicionar e a K, a estimativa da probabilidade condicional não se altera, é possível afirmar que o novo valor $P'(h)$ que o investigador estará disposto a apostar será $P\ (h/e)$.

Ao mesmo tempo, é forçoso admitir que indícios científicos dificilmente implicam consequências dedutivas que condicionalizam, de qualquer forma, estados futuros de crença. Diante desse fato, Howson e Urbach concluem que o princípio de condicionalização apenas descreve a forma como um investigador ideal, que tenha plena consciência do impacto de e em h, dado K, e que tenha expressado tal impacto como $P\ (h/e)$, reagirá, uma vez que tenha efetivamente tomado consciência da ocorrência de e. Embora nenhum indivíduo corresponda ao investigador ideal descrito, isso não impede a utilização do enfoque bayesiano, cuja correta aplicação, asseveram Howson e Urbach, exige o conhecimento de relações dedutivas cuja extensão nenhum ser humano domina em sua inteireza.

A condicionalização bayesiana pode ser entendida, ainda, como um caso particular da chamada condicionalização de Jeffrey. Assim como a condicionalização bayesiana, a condicionalização de Jeffrey também é valida apenas caso a modificação da função de distribuição de probabilidades de P para P' não altere os valores das probabilidades condicionais passíveis de serem obtidas a partir de P. Também aqui a exigência pode parecer excessiva, mas, na verdade, ela apenas significa que o impacto do indício e sobre a hipótese h é expresso de forma perfeita por meio da probabilidade condicional $P\ (h/e)$. Dessa forma, uma vez verificado e, a passagem de P para $P´$ se dá por meio da probabilidade condicional, sem que haja qualquer outro fator em jogo.

A condicionalização de Jeffrey pode ser explicada a partir do seguinte exemplo: imagine uma pessoa a observar um pedaço de tecido com o auxílio de uma vela. Antes da observação, essa pessoa tem graus de crença $P(b), P(g)$ e $P(v)$ de que o tecido é, respectivamente, azul, verde e violeta. Após a observação, esses graus de crença são modificados para $P´(b), P´(g)$ e $P´(v)$. A modificação ocorre de forma tal que as diferenças $P(b)\ - P´(b),\ P(g)\ - P´(g)$ e $P(v)\ - P´(v)$ são distintas e diferentes de zero. Seria possível encontrar uma proposição e que permitiria mediar a modificação de P para $P´$ através da condicionalização bayesiana? Em outras palavras, seria possível encontrar uma proposição e de tal forma que $P´(b)\ = P(b/e), P´(g)\ = P(g/e)$ e $P´(v)\ = P(v/e)$? Ressalte-se que, no exemplo hipotético de Jeffrey, a passagem de P para $P´$ se dá de forma arbitrária, não condicionalizada. A pergunta é, portanto, se seria possível obter uma proposição e a partir da qual seria possível obter as mesmas alterações nas probabilidades originais, mas agora por meio da condicionalização.

Embora nem sempre seja possível encontrar uma proposição e que cumpra o papel descrito acima, Jeffrey demonstra que é possível conhecer a probabilidade posterior $P´(c)$ de qualquer proposição c a partir da probabilidade de outra proposição, por exemplo d, cuja probabilidade se modificou de $P(d)$ para $P´(d)$ de forma não condicionada. Assim, temos que $P´(c)\ = P(c/d)\ \times P´(d)\ +\ P(c/{\sim}d)\ \times P´({\sim}d)$. Howson e Urbach (1989, p. 107) enfatizam que a regra de condicionalização de Jeffrey constitui uma expressão genérica da qual a condicionalização bayesiana pode ser considerada um caso particular. A vantagem da regra de Jeffrey é que ela permite recuperar a distribuição de probabilidades original, o que nem sempre é possível por meio do Teorema de Bayes. Em outras palavras, é possível alterar a probabilidade $P(d)$ de forma não condicionada para $P´(d)$ e, em seguida, novamente para $P(d)$, e obter a distribuição original P para todas as proposições por meio de aplicações sucessivas da regra acima.

Tanto quanto a condicionalização bayesiana, também a condicionalização de Jeffrey foi o alvo de tentativas de estabelecimento de cadernos de apostas holandeses (Howson e Urbach, 1989, pg. 109). Assim como no caso bayesiano, a possibilidade do estabelecimento de uma estratégia de aposta que levará a uma perda certa tampouco prova que a condicionalização de Jeffrey é a única forma consistente de atualização de probabilidades.

A conclusão relevante para o bayesianismo que se pode extrair da discussão acima é que esse enfoque não é aplicável apenas às situações idealizadas em que um pesquisador atualiza seus graus de crença por meio do Teorema de Bayes. Embora o bayesianismo apresente um componente normativo e, assim, tenha a pretensão de indicar como os investigadores devem se comportar do ponto de vista do método científico, ele também permite analisar situações reais. A condição para que o bayesianismo possa auxiliar a analisar o raciocínio científico é que a atualização de probabilidades se dê de forma consistente, mesmo que isso não se dê através da condicionalização bayesiana ou outra. Com isso, Howson e Urbach querem responder a eventuais críticas de que seu enfoque apenas é aplicável em contextos altamente idealizados, não se prestando para compreender a realidade do dia a dia da prática científica.

CONCLUSÃO

Neste capítulo, foram apresentados os fundamentos do bayesianismo para explicar o raciocínio científico. Conforme se viu, esse enfoque está baseado na aplicação do Teorema de Bayes para atualizar o grau de crença na hipótese *h* diante da ocorrência do indício *e*. O enfoque exige, ainda, que a atribuição das probabilidades prévias requeridas pelo Teorema seja feita com base na teoria subjetivista das probabilidades, bem como que tal atribuição respeite os axiomas do cálculo das probabilidades para ser consistente. Ao apresentar esse enfoque, os bayesianos procuram oferecer uma solução definitiva para um antigo problema filosófico, relativo à justificação da indução. Para eles e, mais especificamente, para Howson e Urbach, em cujo livro a exposição precedente se baseou, a aplicação do Teorema de Bayes constituiria a verdadeira "lógica" justificadora do raciocínio indutivo, cujo papel pode ser equiparado ao da lógica clássica para o raciocínio dedutivo. Além disso, ao apresentar o bayesianismo, outro objetivo central de Howson e Urbach, conforme se viu nas seções precedentes, é mostrar que o bayesianismo não apenas dá conta de uma série de aspectos importantes da atividade científica, como o faz de maneira superior a enfoques alternativos, especialmente o falsificacionsimo popperiano.

Desde suas formulações iniciais, o bayesianismo foi alvo de críticas diversas, especialmente no que diz respeito ao exacerbado subjetivismo envolvido na atribuição das probabilidades prévias. Essas críticas foram rebatidas e, como Howson insiste em afirmar na última edição de *Scientific Reasoning,* grande parte delas deixou de ser relevante após as (definitivas) explicações bayesianas. Mesmo assim, o bayesianismo está longe de lograr avaliações consensuais quanto aos seus méritos, não obstante sua aplicação tenha se expandido, nas últimas décadas, para várias áreas do conhecimento. É por essa razão que, no próximo capítulo, serão examinadas com vagar as principais críticas ao enfoque bayesiano e as respectivas respostas fornecidas pelos seus defensores.

2. CRÍTICAS AO ENFOQUE BAYESIANO DO RACIOCÍNIO CIENTÍFICO RESPONDIDAS POR HOWSON & URBACH

Neste capítulo, serão apresentadas críticas ao enfoque bayesiano. Serão analisadas não apenas as críticas que Howson e Urbach consideram as mais relevantes do ponto de vista filosófico e que, nessa condição, constam da terceira e última edição de *Scientific Reasoning – the Bayesian approach*, mas, igualmente, as objeções que foram respondidas nas duas edições anteriores daquele livro, tendo sido retiradas da última edição, de 2006.

Na primeira parte do capítulo, serão apresentadas a questão do alegado subjetivismo inerente ao enfoque bayesiano e o problema do indício antigo, que constituem as objeções conservadas por Howson e Urbach na edição de 2006. Na segunda parte, serão apresentadas as críticas desenvolvidas nas edições anteriores, as quais são devidas, em grande parte, a Popper e seus seguidores. Tanto quanto na apresentação do bayesianismo tratada do capítulo anterior, também ao rebater as críticas que lhes foram endereçadas persiste a intenção de Howson e Urbach de demonstrar a superioridade do bayesianismo com relação a enfoques alternativos, em particular o falsificacionismo.

O SUBJETIVISMO

Possivelmente a crítica mais incisiva ao enfoque bayesiano é a de subjetivismo na atribuição das probabilidades prévias. Conforme se viu, o bayesianismo não impõe qualquer limitação à atribuição de valores para as probabilidades prévias no Teorema de Bayes (salvo, obviamente, que sejam respeitados os axiomas do cálculo de probabilidades). Tal subjetivismo seria alegadamente incompatível com o propósito bayesiano de fornecer uma interpretação adequada para o raciocínio indutivo. Uma vez que os valores das probabilidades prévias são, necessariamente, subjetivos, seria impossível compatibilizar o enfoque bayesiano com o caráter objetivo que se supõe ser uma característica essencial do raciocínio científico.

Ao responder à objeção de que o bayesianismo dá excessivo espaço a aspectos subjetivos, Howson e Urbach desenvolvem uma estratégia concebida de forma progressiva. Inicialmente, os autores indicam que é impossível fugir a certo grau de subjetivismo, independentemente do enfoque adotado para analisar o raciocínio científico. Em seguida, os autores atacarão a ideia mesma de que é possível se chegar a probabilidades prévias objetivas. Assim, eles indicarão que: 1) a busca por probabilidades prévias objetivas é uma quimera, pois é praticamente impossível que dois cientistas coincidam em sua avaliação de uma hipótese; 2) à medida que os indícios se acumulam, o aspecto subjetivo da inferência baseada no Teorema de Bayes tende a se diluir; e 3) os critérios que supostamente servem para constranger a liberdade do analista na atribuição de probabilidades prévias e torná-las objetivas, como o "princípio da indiferença" ou a noção de simplicidade, não são universalmente válidos.

Nas primeiras páginas do capítulo final de *Scientific Reasoning* (2006), dois pontos preliminares são ressaltados: (1) certo grau de subjetivismo é inerente a qualquer avaliação científica. O mérito do bayesianismo é, precisamente, o de indicar explicitamente onde está o aspecto subjetivo, em lugar de "escondê-lo debaixo do tapete"; (2) o enfoque bayesiano oferece uma lógica objetiva para o raciocínio indutivo. As probabilidades prévias equivaleriam às premissas do raciocínio dedutivo, o Teorema de Bayes seria o mecanismo apto a gerar conclusões válidas, as quais, por sua vez, corresponderiam aos valores das probabilidades posteriores (Howson &Urbach, p. 264, 2006).

Howson e Urbach lembram que, na prática, dois cientistas dificilmente coincidem na avaliação de determinada hipótese, especialmente nos estágios iniciais de investigação. Os autores chegam a afirmar (2006, p. 238) que a ideia de que teorias científicas ou hipóteses apresentam valores de probabilidade prévia passíveis de serem determinados objetivamente, embora atrativa, se choca com a avaliação consensual entre os cientistas de que é impossível estabelecer tais valores. A busca por um valor "objetivo" para a probabilidade prévia de uma hipótese seria, assim, uma quimera[20].

É possível, por outro lado, formular uma defesa da objetividade do enfoque bayesiano por meio do conceito matemático de limite. Dois cientistas podem, no limite, convergir em suas probabilidades posteriores, independentemente dos valores prévios atribuídos individualmente à hipótese em exame. O problema é, porém, que tal defesa do caráter objetivo do enfoque bayesiano só se mostra efetiva no limite, isto é, no caso de uma quantidade infinita de indícios, não sendo aplicável para uma quantidade finita.

---

[20] Entre os bayesianos que defenderam a adoção de probabilidades prévias objetivas está Jaynes que, em *Probability Theory: the Logic of Science*, levou adiante os trabalhos de Jeffreys sobre esse assunto. Jaynes defendia que as distribuições prévias de probabilidades deveriam ser objetivas no sentido de serem completamente independentes da personalidade do investigador. O problema é que os critérios sugeridos para alcançar esse fim não são, como sustenta Howson, universalmente válidos, o que põe em risco toda a proposta dos bayesianos objetivistas.

Por essa razão, Howson e Urbach reputam insuficiente tal tentativa de defesa do bayesianismo, tanto do ponto de vista normativo quanto explicativo (Howson & Urbach, 2006, p. 238).

Frente à dificuldade levantada, Howson e Urbach apresentam um novo argumento que, embora similar ao anterior, se mostra capaz de defender o bayesianismo das acusações de falta de objetividade e evitar as limitações levantadas pela aplicação do conceito de limite. O argumento se baseia na constatação de que as opiniões científicas frequentemente convergem rapidamente para um valor único, após o acúmulo de relativamente poucos indícios. Nessa situação, o Teorema de Bayes proporciona o arcabouço explicativo para o fenômeno, inclusive ao oferecer explicações convincentes para as peculiaridades que possam surgir em cada caso específico. Dois exemplos são apresentados (Howson & Urbach, 2006, pp. 239-243): 1) a estimativa da média de uma população que segue o padrão normal de distribuição normal; e 2) a estimativa de uma proporção no caso de uma distribuição binomial.

O primeiro exemplo diz respeito a uma população cuja distribuição já se sabe, de antemão, seguir o padrão normal, com desvio padrão σ. Em termos práticos, essa situação poderia ser exemplificada pela realização de medidas de determinada grandeza física por meio de um instrumento de medição qualquer. Utilizado repetidamente e em condições similares, o instrumento forneceria um conjunto de medidas cuja variação se aproxima de uma curva de distribuição normal (curva em forma de sino invertido). O registro das medições poderia ser considerado equivalente a uma amostra obtida ao acaso a partir de uma população de distribuição normal cuja média é desconhecida e cujo desvio padrão foi estabelecido previamente.

Howson e Urbach mostram que, à medida que a amostra é ampliada, sua média se aproxima da média da população em exame. Isso significa que, com o aumento da amostra, a influência da distribuição prévia de probabilidades da média– e, dessa forma, do aspecto subjetivo da inferência – diminui de forma progressiva, vindo, no fim das contas, a tornar-se desprezível. Dois cientistas cujas probabilidades prévias, expressas em curvas de distribuição normal, sejam distintas, chegarão a distribuições posteriores coincidentes após o acúmulo suficiente de indícios, o que ocorre com relativa rapidez. O ponto crucial enfatizado pelos autores (2006, p. 240) é, portanto, que na realização de uma inferência com base na aplicação do Teorema de Bayes, os dados objetivos contidos na amostra se tornam o fator preeminente de forma relativamente rápida. A única restrição, aqui, é que, em se tratado de amostras suficientemente grandes, a atribuição das probabilidades prévias não deve ser extrema para que a convergência ocorra.

O enfraquecimento do aspecto subjetivo da inferência e a convergência de opiniões se dão de maneira relativamente rápida. No exemplo aduzido pelos autores, dois cientistas com probabilidades prévias acentuadamente divergentes chegam aos mesmos valores para média e desvio padrão da população com uma amostra integrada por cem elementos (2006, p. 241). Na verdade, bastaria que a amostra fosse composta por vinte elementos para que as distribuições posteriores de probabilidade já estivessem bastante próximas. Dessa forma, apesar de a análise bayesiana partir de distribuições prévias de probabilidade subjetivas, a conclusão extraída por meio do mecanismo indutivo baseado no Teorema de Bayes é objetiva e está ancorada nos indícios reunidos.

O segundo exemplo aduzido por Howson e Urbach tem por finalidade demonstrar que a diluição progressiva dos aspectos subjetivos da inferência bayesiana também se verifica em situações mais gerais do que o de uma população com distribuição normal. Desta vez, as situações práticas referem-se, por exemplo, à estimativa da proporção de bolas vermelhas em uma urna que contenha bolas vermelhas e verdes ou, ainda, a probabilidade de uma moeda exibir "cara" ao ser lançada. Os indícios são recolhidos por meio da retirada sucessiva, com reposição, de bolas da urna ou, na segunda situação, pela anotação, a cada lançamento da moeda, do resultado "cara" ou "coroa". O importante, aqui, é que só há dois resultados possíveis para cada experimento (vermelho x verde; cara x coroa), cujas probabilidades são, portanto, θ e $1-\theta$. Quando isso ocorre, estamos diante de um fenômeno que, em estatística, é conhecido como "Processo de Bernoulli" e no qual os parâmetros ou proporções binomiais são, precisamente, θ e $1-\theta$.

Também aqui é assumida uma distribuição hipotética para curva de probabilidades prévias, conhecida como "distribuição beta". Toma-se uma amostra de n tentativas derivadas de um Processo de Bernoulli, que produz s sucessos e f fracassos $(s+f=n)$. Nessas condições, a média da distribuição posterior de probabilidades tente a $\frac{s}{n}$, na medida em que o número de observações cresce em direção ao infinito; a variação da distribuição posterior de probabilidades tende, por sua vez, a zero. Isso significa que, assim como no exemplo anterior, a influência da distribuição prévia de probabilidades com relação à distribuição posterior diminui proporcionalmente com o aumento do tamanho da amostra.

Com o propósito de atestar a validade general dessas conclusões acerca da inferência bayesiana, em uma seção posterior de *Scientific Reasoning* (2006, p. 245), Howson e Urbach recorrem a demonstrações matemáticas para explicar o chamado "Princípio da Estimativa Estável" e concluem que o referido princípio garante que a independência das estimativas bayesianas com relação às distribuições prévias de probabilidade não está restrita a tipos específicos de distribuição (normal e Bernoulli). O princípio também garante que, partindo-se de uma amostra suficientemente grande, não é sequer necessário descrever a função de probabilidades prévias de forma exata para realizar uma inferência acurada e precisa baseada no Teorema de Bayes.

Em outra seção do livro, Howson e Urbach procuram defender o enfoque bayesiano da acusação de subjetivismo por meio de uma linha argumentativa complementar à exposta acima, fundada na pouca influência das probabilidades prévias na inferência baseada no Teorema de Bayes. Assim, após concluir que a influência dos aspectos subjetivos na inferência bayesiana é muito menor do que se imagina inicialmente, os autores querem mostrar que qualquer tentativa de estabelecer parâmetros "objetivos" para a atribuição das probabilidades prévias é artificial e, em última análise, ilusória. Assim, parâmetros que, na prática, deveriam funcionar como restrições à liberdade do cientista de atribuir valores à probabilidade prévia de sua hipótese, como o "princípio da indiferença" e a exigência de "simplicidade" serão, um a um, postos de lado.

O "princípio da indiferença" estabelece que, em uma situação em que inexiste informação de fundo, devem-se atribuir probabilidades iguais a partes iguais do universo de possibilidades. Trata-se de um princípio intuitivamente válido, segundo o qual deve haver uma relação de simetria entre os membros de uma partição ao se atribuírem probabilidades *a priori*. Assim, no caso de um dado sobre o qual não se sabe se é viciado ou não, é natural atribuir-se a probabilidade de $\frac{1}{6}$ a cada dos resultados possíveis. Esse seria o valor da probabilidade prévia a ser atribuída a cada um dos resultados, antes que o acúmulo de indícios resultantes de lançamentos sucessivos pudesse indicar se o dado está ou não viciado. O princípio constitui, ainda, um dos fundamentos de demonstrações levadas a cabo por Bernoulli e por Bayes para mostrar que, em uma amostra suficientemente grande (n tende ao infinito), a probabilidade *p* de um evento *r* qualquer é aproximadamente igual a sua frequência relativa, isto é, $\frac{r}{n}$.

Apesar de seu caráter intuitivo, o princípio da indiferença pode levar a inconsistências que colocam em dúvida sua validade em todas as situações. Com o auxílio da lógica formal (Howson e Urbach, 2006, p. 270), é possível mostrar que a aplicação do princípio em determinadas linguagens tem, como consequência, a atribuição de probabilidades diferentes a um mesmo fenômeno, o que obviamente é inconsistente.

Em um livro posterior (Howson, 2011, p. 121), Howson utiliza o seguinte exemplo para mostrar que há boas razões para não se tomar o princípio da indiferença como absoluto: a partição mais elementar do universo de possibilidades para um sistema quântico de um par de bósons, cada um dos quais pode estar no estado A ou no estado B deveria, intuitivamente e de acordo com o princípio da indiferença, ser composto por quatro possibilidades, isto é, (A,A), (A,B), (B,A), (B,B). Em função de determinadas particularidades da física quântica, porém, o universo de possibilidades tem apenas três membros, cada um dos quais tem probabilidade de $\frac{1}{3}$: “ambos A”, “um A e um B” e “ambos B”. A física quântica simplesmente não permite considerar (A,B) e (B,A) como possibilidades distintas, o que levanta o problema, do ponto de vista científico, de que as partículas não são suscetíveis de individuação, muito embora constituam um conjunto plural.

Mesmo que se tomem partições mais elementares do universo de resultados possíveis, é possível, ainda assim, chegar a atribuições de probabilidade diferentes, dependendo de como tais partições são expressas, o que representa uma contradição (Howson e Urbach, 2006, p. 272). É verdade, por outro lado, que historicamente o bayesianismo baseou-se, em suas origens, no princípio da indiferença e que bayesianos objetivistas, como Jeffreys, tentaram a todo custo encontrar maneiras de reconciliar o enfoque bayesiano com as implicações problemáticas daquele princípio, convencidos de que este conferiria base sólida e objetiva para a indução probabilística. Howson e Urbach indicam, porém, que tais tentativas, não obstante terem chegado a resultados interessantes, não estão isentas de problemas, um dos quais é que levariam a funções de probabilidades prévias impróprias, isto é, que violariam os axiomas do cálculo de probabilidades.

É verdade que seria possível encontrar fórmulas para resolver as inconveniências criadas pelas probabilidades prévias impróprias. Não obstante, ficaria em aberto a questão de por que se deveriam adotar as regras propostas por Jeffreys. O enfoque do bayesianismo proposto por Howson e Urbach toma, como definição de probabilidade, a noção do valor que um apostador está disposto a empenhar em uma aposta justa (*fair betting odds*). Em um primeiro momento, acreditou-se que o princípio da indiferença era consistente com essa noção, mas exame mais detalhado evidenciou as inconsistências dele decorrentes, conforme se viu. No caso das regras propostas por Jeffreys, não parece haver forma de justificar sua inclusão no arcabouço teórico do bayesianismo, tendo em vista a referida noção de probabilidade adotada (Howson e Urbach, 2006, pp.275-76).

A atratividade exercida pelo princípio da indiferença se explica, em grande parte, por seu caráter intuitivo, uma vez que ele determina que cada elemento da partição do universo de possibilidades deve ser considerado igualmente provável *a priori*, como no exemplo do dado acima. Argumenta-se que haveria, dessa maneira, uma garantia de que a inferência seria objetiva. O fato é, porém, que há inúmeras distribuições de probabilidades para os elementos do universo de possibilidades. Toda e qualquer distribuição prévia carrega consigo algum tipo de implicação que vai além do que é expresso pelos indícios. A conclusão é, portanto, que não existe uma distribuição *a priori* de probabilidades, como se pretende que seja aquela feita com base no princípio da indiferença, que permita aos indícios “falarem por si mesmos”.

Outro critério discutido por Howson e Urbach em seu intento de mostrar a impossibilidade da determinação de probabilidades prévias objetivas diz respeito ao critério de simplicidade. Do ponto de vista bayesiano, o critério da simplicidade implica que se deve atribuir probabilidade prévia mais alta às hipóteses ou teorias mais simples. Apesar de parecer intuitivamente justificado, o critério da simplicidade apresenta problemas. Não há, por exemplo, um entendimento unívoco acerca do que significa afirmar que uma hipótese ou teoria é mais simples do que outra. O conceito de simplicidade se refere a elementos distintos em contextos diferentes. A depender de como é caracterizada, determinada teoria pode ser considerada mais ou menos simples. Um primeiro problema é, assim, que não existe uma definição universalmente válida para o conceito de simplicidade, o qual pode se mostrar de difícil aplicação no contexto científico.

É verdade, por outro lado, que é possível adotar, como definição precisa para simplicidade, o conceito atribuído a Guilherme de Occam, segundo o qual "as entidades não devem ser multiplicadas além do necessário, a natureza é por si econômica e não se multiplica em vão"[21]. Howson e Urbach interpretam as palavras de Occam de modo que hipóteses e/ou teorias mais simples seriam aquelas com o menor número de parâmetros independentes ajustáveis. Tal definição, reconhecem Howson e Urbach, exerce um apelo imediato e evidente junto à comunidade científica, que intuitivamente acredita ser razoável que uma hipótese com menor número de parâmetros independentes, *ceteris paribus*, tenha mais chances de ser verdadeira e, portanto, deva ter probabilidade prévia mais alta que outra.

O contra-argumento de Howson e Urbach é que, embora inexista qualquer problema técnico com o critério de simplicidade em si (em seu afã objetivista, Jeffreys chega a utilizar a expressão "postulado da simplicidade"), é pouco provável que esse critério possa ser de qualquer utilidade na avaliação de hipóteses sem que se levem em consideração as informações de fundo e as especificidades do caso concreto. Mesmo a aversão professada por cientistas a hipóteses com muitos parâmetros ajustáveis só pode ser levada a sério à luz dos méritos intrínsecos do caso em exame. A conclusão de Howson e Urbach é que, na verdade, o critério determinante na comparação de hipóteses não é a noção de simplicidade, mas, sim, a de plausibilidade. É possível imaginar inúmeras situações em que uma hipótese mais simples seria rejeitada por outra dotada de maior plausibilidade, embora mais complexa. Na área econômica, por exemplo, na qual há, segundo os autores, um senso compartilhado da existência de vários fatores causais independentes, qualquer modelo com poucas variáveis costuma ser subliminarmente descartado (Howson e Urbach, 2006, p. 291).

A posição de Howson e Urbach com relação à noção de simplicidade não implica, em absoluto, que haja qualquer coisa de errado em adotar hipóteses mais simples em situações concretas. Em igualdade de condições, hipóteses com menos parâmetros costumam ser mais bem confirmadas pelos indícios recolhidos. Do ponto de vista bayesiano, porém, "igualdade de condições" significa que as hipóteses em comparação já apresentam probabilidades prévias equivalentes, o que é distinto de adotar o conceito de simplicidade como critério absoluto para atribuir probabilidades prévias. Assim, Howson e Urbach consideram que o critério de simplicidade é inadequado para a atribuição de probabilidades prévias a hipóteses. O critério adequado, para os autores, é o de plausibilidade, cuja utilização nem sempre levará aos mesmos valores das atribuições fundadas *a priori* na ideia de simplicidade.

[21] Para maiores considerações sobre a noção de simplicidade e sua relação com o bayesianismo, ver Swinburne (2001).

Em resumo, Howson e Urbach acreditam que a tentativa de estabelecer probabilidades prévias objetivas constitui um empreendimento desnecessário e errôneo. A busca por probabilidades prévias objetivas levanta problemas cuja solução, para eles (Howson e Urbach, 2006, pp.297), só se torna possível ao se recorrer a expedientes baseados em decisões subjetivas, que é exatamente o que se pretendia evitar inicialmente.

Cientistas costumam avaliar diferentemente hipóteses, mesmo em situações em que a informação de fundo compartilhada seja a mesma. Não há nada errado com isso. Tentar fazer com que haja uniformidade de opinião, como tentam os bayesianos objetivistas, pode se mostrar deletério do ponto de vista do progresso científico, especialmente porque significaria abrir mão de informações para chegar às probabilidades prévias uniformes e "objetivas".

Howson e Urbach acreditam, portanto, que os axiomas da probabilidade constituem parâmetros suficientes para garantir a consistência da inferência bayesiana, que nada deve à lógica dedutiva do ponto de vista da objetividade de suas conclusões. A forma mais natural de interpretar o mecanismo bayesiano é de que se trata de um conjunto de regras que permite se chegar a conclusões probabilísticas a partir de premissas probabilísticas. Tais conclusões serão tão acuradas quanto o forem as premissas de que se parte. A objetividade da inferência se explica a partir da validade objetiva com que as conclusões são inferidas das premissas, o que é garantido pela aplicação do Teorema de Bayes (Howson e Urbach, 2006, pp 296-297).

O PROBLEMA DO INDÍCIO ANTIGO

Ao iniciar sua discussão do Problema do Indício Antigo (*The Old Evidence Problem)*, Howson e Urbach fazem questão de frisar tratar-se de uma objeção que não mereceria discussão séria e que é mencionada apenas por que "há pessoas que objetam a qualquer teoria e o enfoque bayesiano não é exceção" (Howson & Urbach, 2006, p. 298). Não obstante, os autores reconhecem que devem tratar dessa suposta objeção, uma vez que ela é, com frequência, apresentada como a crítica mais séria ao emprego da interpretação subjetivista das probabilidades na tentativa de avaliar o progresso científico.

O problema pode ser expresso da seguinte forma: como se sabe, o enfoque bayesiano constitui uma forma de raciocínio baseada na maneira como o acúmulo de indícios modifica as probabilidades subjetivas do pesquisador. Uma situação particular que surge, nesse contexto, diz respeito ao impacto em uma hipótese h, relativa ao indício *e*, que está baseada em indícios recolhidos *antes* da apresentação da hipótese. O exemplo clássico são as anomalias registradas na órbita de Mercúrio em meados do século XIX, que são consideradas como indícios de confirmação da Teoria Geral da Relatividade, de Einstein, descoberta apenas em 1915. Críticos alegam que o enfoque bayesiano não pode, em princípio, explicar essa situação, uma vez que o indício *e*, por já ser conhecido, implica que $P(e) = 1$. Nessas condições, a probabilidade posterior de *h*, isto é, $P(h/e) = P(h)$, o que significa, contra o que se poderia esperar, que os indícios já conhecidos não exercem qualquer efeito para modificar a probabilidade prévia de *h*.

A defesa de Howson e Urbach está baseada na convicção de que a questão do indício antigo não representa qualquer ameaça ao arcabouço teórico bayesiano, mas representa, tão somente, a incapacidade de utilizá-lo corretamente. Isso ocorre por que nenhum indício, tomado isoladamente, é capaz de confirmar ou infirmar uma hipótese. O efeito de um indício com relação a determinada hipótese somente pode ser avaliado à luz das informações de fundo à disposição do pesquisador.

Confirmar ou infirmar constitui, assim, uma relação que envolve três elementos: o indício, a hipótese e um conjunto *k* de informações de fundo. A capacidade de *e* de confirmar ou infirmar é medida pela sua influência sobre a credibilidade de *h*, dada a existência de *k*. Diante dessas considerações, a única forma possível para avaliar, no caso do indício antigo, o impacto de *e* sobre *h* é garantir que *e* não esteja contido em *k*. Deve-se, portanto, descontar *e* de *k* para que a aplicação do Teorema de Bayes possa ser feita de maneira correta.

Em lugar de se tratar de um problema insolúvel, o problema do indício antigo torna-se, à luz das considerações precedentes, apenas uma questão de se encontrar a melhor forma de descontar o indício antigo do conjunto de informações de fundo. Howson e Urbach, a essa altura, reconhecem que não se trata de uma tarefa simples, mas asseguram, com base na literatura especializada, ser perfeitamente possível estabelecer funções de probabilidade que expressem a eliminação do indício *e* de *k*.

Especificamente com relação ao exemplo concreto relativo às anomalias na órbita de Mercúrio e a Teoria Geral da Relatividade, Howson e Urbach partem da seguinte versão do Teorema de Bayes para responder à crítica levantada e demonstrar a validade do enfoque proposto (Howson & Urbach, 2006, p. 300):

$$P(h/e) = \frac{p}{p + \frac{P(e/\sim h)}{P(e/h)} \times (1-p)}$$

Ao aplicar o Teorema de Bayes, tem-se que h representa a Teoria Geral da Relatividade; e constitui o conjunto de indícios relativos às anomalias na órbita de Mercúrio; *p* corresponde à probabilidade prévia de h, isto é, a crença do investigador na hipótese descontado qualquer impacto atribuível a e. Como e constitui um indício dedutível a partir da Teoria Geral da Relatividade, tem-se, com base nos axiomas do cálculo de probabilidades, que $P(e/h) = 1$. Com isso, a versão acima do Teorema de Bayes é simplificada para:

$$P\,(h/e) = \frac{p}{p + P(e/\sim h) \times (1-p)}$$

Nessa nova equação, faz-se necessário, para calcular o impacto de *e* com relação a h, estabelecer os valores de p e $P(e/\sim h)$. O valor de $P(e/\sim h)$ corresponde à probabilidade de *e* ser explicável pela única hipótese alternativa a h, no caso, a Teoria Clássica da Gravitação de Newton (na verdade isso é uma simplificação, pois $P(e/\sim h)$ corresponde a uma constante multiplicada pela somatória de todos os produtos $P(e/h_i) \times P(h_i)$, onde h_i corresponde às hipóteses alternativas a *h*). Como a Teoria Clássica da Gravitação não explica de maneira satisfatória as anomalias na órbita de Mercúrio, é natural assumir que $P(e/\sim h)$ será um número bem pequeno, por exemplo, ε. A probabilidade prévia *p*, cuja avaliação é feita com base nas informações de fundo descontando-se *e*, por sua vez, será um número também pequeno, mas não desprezível. Howson e Urbach não explicam porque não consideram o valor de *p* desprezível. Provavelmente isso ocorre em função da existência de indícios de confirmação que justificam que seja atribuída a *p* uma probabilidade prévia diferente de zero. De qualquer forma, a avaliação que fazem é condizente com o enfoque subjetivista na atribuição das probabilidades prévias. É razoável imaginar que, no início do século XX, havia indícios suficientemente convincentes a apoiar a Teoria de Einstein, mesmo sem levar em consideração a questão da órbita de Mercúrio. Com isso, teremos que a equação acima resultará na divisão de um número pequeno por outro apenas ligeiramente maior. O resultado dessa divisão será próximo a 1, o que evidenciaria o poder de confirmação de *e* com relação a h, mesmo se tratando de um indício antigo (Howson & Urbach, 2006, pp. 300-301).

Com o afastamento do problema do indício antigo, Howson e Urbach concluem a última versão de *Scientific Reasoning.* Embora considerem que o enfoque que apresentam está longe de ter sido desenvolvido de forma exaustiva, os autores se mostram confiantes de que seus continuadores lograrão aperfeiçoar o modelo bayesiano. As palavras finais de Howson e Urbach quanto às críticas relativas ao excessivo subjetivismo na atribuição das probabilidades prévias são para reiterar que tais críticas são focadas em questões exógenas ao modelo bayesiano. Nas seções seguintes, serão apresentadas outras críticas ao modelo bayesiano também rebatidas por Howson e Urbach, mas extraídas das duas edições anteriores da obra em exame.

O BAYESIANISMO FAVORECERIA HIPÓTESES FRACAS

Popper foi o primeiro a criticar o bayesianismo com base no argumento de que este favoreceria hipóteses logicamente fracas. Na *Lógica da Pesquisa Científica*, Popper afirma categoricamente que "os cientistas devem escolher entre alta probabilidade e alto conteúdo informativo, uma vez que, por razões lógicas, não é possível ter ambos" (Popper, 1959, p. 363). Para Popper, tentativas de explicar as inferências indutivas com base em modificações nas probabilidades prévias de hipóteses, tendo em vista acúmulo de indícios, seriam, necessariamente, enviesadas em favor de hipóteses logicamente mais fracas, ou seja, hipóteses que explicam pouco.

Há uma variante do argumento popperiano, que foi apresentada pela primeira vez por Glymour (Glymour, 1980, pp. 84-85). Segundo a versão de Glymour do argumento original de Popper, os indícios observáveis das teorias científicas seriam pelo menos tão prováveis quanto as próprias teorias. Por essa razão, haveria um vínculo entre crença e capacidade explicativa: quanto maior for a capacidade explicativa de uma hipótese – e, portanto, a quantidade de indícios por ela implicados–, maior deverá ser a crença na mesma, e vice-versa. Isso significa que, para Glymour, maior capacidade explicativa deve ser necessariamente acompanhada por maior probabilidade prévia.

A resposta à crítica de Popper é que nem os axiomas do cálculo de probabilidades, nem a lógica formal impedem a atribuição de probabilidade alta, mesmo o valor 1, a uma hipótese logicamente forte, desde que não se trate de uma contradição. A associação feita por Popper entre força lógica de uma sentença e baixa probabilidade é, para Howson e Urbach, simplesmente incorreta (Howson e Urbach, 1993, p. 390). Popper acredita que hipóteses mais gerais, ao explicar um grande número de fenômenos, terão necessariamente probabilidade alta. No contexto da atividade científica, porém, nada impede que determinada hipótese esteja restrita a fenômenos específicos, seja logicamente forte e tenha, portanto, probabilidade alta.

Com relação à versão de Glymour do argumento de Popper, a resposta bayesiana é de que, embora exista um vínculo entre crença e capacidade explicativa, esse vínculo não é imediato nem tão direto quanto o pretendido por aquele autor. Os bayesianos reconhecem que uma hipótese explicativa forte tenderá a ser objeto de crença e, portanto, tenderá a exibir probabilidade alta, tendo em vista os indícios de confirmação acumulados. Trata-se apenas de uma tendência que deve ser comprovada pelo acúmulo de indícios. Não existe qualquer obrigatoriedade nisso, pois nada impede que o acúmulo de indícios leve a que a hipótese explicativa forte seja desacreditada. Além disso, os bayesianos não aceitam que um aumento no grau de crença seja necessariamente acompanhado por um aumento no poder explicativo da hipótese.

O ponto de contato entre Glymour e Popper é que ambos, embora partam de premissas distintas, desconfiam da capacidade de o bayesianismo apresentar uma explicação coerente para hipóteses de grande poder explicativo. Para Popper, as hipóteses com grande capacidade explicativa, por serem logicamente fortes, devem ter probabilidade baixa. Isso se chocaria com o raciocínio bayesiano, em que tal probabilidade deveria tornar-se alta apenas após o acúmulo contínuo de indícios de confirmação. Para Glymour, por sua vez, as hipóteses com grande capacidade explicativa devem, necessariamente, ter probabilidade alta, ao passo que o bayesianismo afirma que, embora essa seja uma tendência, ela só acontecerá após o acúmulo correspondente de indícios.

A PROBABILIDADE PRÉVIA DE HIPÓTESES UNIVERSAIS DEVE SER ZERO

Popper, como se viu, acredita na impossibilidade de uma hipótese ter simultaneamente amplo conteúdo informativo e probabilidade alta. Como consequência disso, ele defende que hipóteses universais devem, tanto por razões lógicas quanto em função das restrições impostas pelo cálculo de probabilidades, ter probabilidade zero (Popper, 1959, apêndices VII e VIII). Howson e Urbach notam que, se o que Popper defende fosse verdade, todo o edifício bayesiano cairia por terra, pois nenhum indício seria capaz de alterar a probabilidade prévia igual a zero da hipótese em questão. Afinal, se $P(h) = 0,$ então $P(h/e) = 0.$

O argumento de Popper é desenvolvido da seguinte forma: tome-se uma linguagem de predicados simples L(A,*n*) por meio da qual é possível avaliar se o atributo A é possuído por quaisquer ou por todos os *n* elementos distintos a_i. Nessa situação, haveria 2^n sentenças que descreveriam as situações possíveis nesse universo. Popper argumenta que em um caso apenas, dentre os 2^n possíveis, seria possível satisfazer a sentença $\forall x A(x)$. À medida que *n* cresce em direção ao infinito, a proporção de situações possíveis a satisfazer a sentença universal torna-se, no limite, zero. Com isso, Popper acredita ficar demonstrado que se deve atribuir zero, necessariamente, à probabilidade de leis universais.

Para responder a essa crítica, Howson e Urbach retomam argumentos apresentados anteriormente, relativos ao "princípio da indiferença". Ao desenvolver seu raciocínio, Popper implicitamente assume que qualquer situação dentre as 2^n possíveis é igualmente provável e apresenta probabilidade prévia igual a $\frac{1}{n}$. Com isso, a probabilidade da única situação a satisfazer a sentença $\forall x A(x)$ é, no limite, zero, à medida que *n* tende ao infinito. O problema é que em nenhum momento Popper justifica a razão de atribuir, com base no princípio da indiferença, probabilidades iguais para cada uma das situações possíveis do conjunto 2^n. Afinal, nenhuma partição do espaço lógico de possibilidades é neutra do ponto de vista epistêmico, nem mesmo a que atribui a mesma probabilidade prévia a cada um dos elementos de um conjunto de resultados possíveis. Além disso, ao atribuir à sentença "não há leis universais" o valor de probabilidade prévia 1, o pesquisador chega a uma situação em que essa probabilidade permanece inalterada, não obstante a existência de fortes indícios de que o fenômeno em questão exemplifica um comportamento conforme determinada lei.

O fato de Popper atribuir às leis universais a probabilidade prévia zero pode ser considerado como uma consequência direta da forma como ele entendia o progresso científico. Para Popper, a história da ciência seria a história de teorias explicativas que foram, paulatinamente, objeto de tentativas de refutação. A partir desse ponto de vista, é possível compreender as razões que, para Popper, justificariam a atribuição de probabilidade zero a qualquer hipótese universal: ela será refutada inevitavelmente.

A pergunta que fazem Howson e Urbach é se, de fato, essa conclusão pessimista seria a lição deixada pela história da ciência. Os autores discordam de Popper por duas razões principais. Primeiro, as falsificações de hipóteses propostas no passado não consistem em nenhuma garantia de que todas as hipóteses que venham a ser apresentadas daqui em diante também serão, necessariamente, refutadas. Segundo, porque não é evidente que hipóteses universais sejam rejeitadas em sua inteireza. Na maioria das vezes, o que os cientistas fazem, diante de indícios que infirmam suas teorias, é seguir trabalhando e promover alterações que, acredita-se, permitirão preservar uma boa parte da teoria original. Além disso, mesmo naquelas situações em que teorias foram descartadas em sua inteireza e sua probabilidade posterior tornou-se zero ou próxima de zero, isso apenas se verificou por que, antes de se tornar desacreditada, aquela teoria era objeto de crença e tinha probabilidade prévia distinta de zero. O que houve, portanto, foi um processo paulatino de infirmação da teoria/hipótese com base nas conclusões obtidas a partir do conjunto de indícios acumulados. A probabilidade posterior igual a zero somente pode ser atribuída *a posteriori,* nunca *a priori,* como quer Popper.

Em suma, a defesa bayesiana contra essa crítica de Popper pode ser resumida da seguinte forma: a) os argumentos lógicos pela atribuição de probabilidade zero para leis universais não são válidos; b) a referida atribuição de probabilidade é incoerente à luz do acúmulo de indícios que confirmam uma hipótese; e c) a história da ciência não oferece evidência de que o valor da probabilidade prévia para leis universais deveria ser zero (Howson e Urbach, 1993, p. 395).

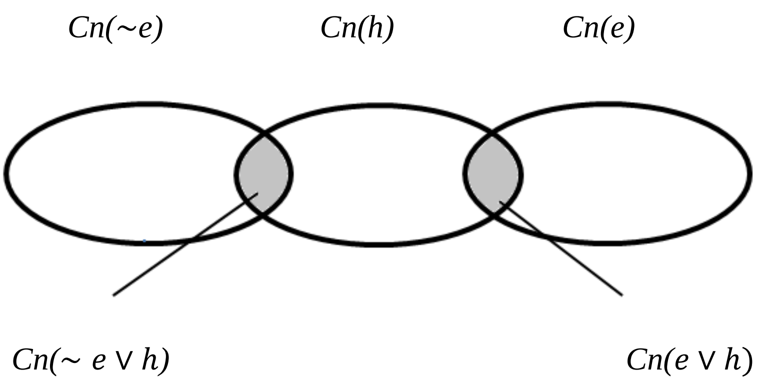
IMPOSSIBILIDADE DA INDUÇÃO PROBABILÍSTICA

Também nesse caso a crítica ao enfoque bayesiano foi articulada por Popper, que, um artigo publicado com D.W. Miller, intitulado *A proof of the impossibility of inductive probability* (1983), defendeu a impossibilidade de realização de qualquer inferência relativa a uma hipótese *h*, tendo em vista a ocorrência do indício *e*. Conforme se viu, o cientista bayesiano acredita que o indício *e* confirma a hipótese *h* se, e somente se, $P\,(h/e)\ > P(h)$. Dos axiomas do cálculo das probabilidades, viu-se, igualmente, ser necessário que $P(h) > 0$ e que $P(e) < 1$.

Popper e Miller argumentam no referido artigo que, uma vez verificadas as condições descritas acima e, adicionalmente, a condição segundo a qual $P(h/e)\ < 1$, então a expressão $\sim e \vee h$ expressa o "excesso indutivo" ou "o conteúdo excessivo" de *h* com relação a *e*, isto é, a parte do conteúdo de *h* que vai além de *e*. Tal excesso de conteúdo de *h* seria, para os autores, infirmado por *e*. A conclusão de Popper e Miller, que é paradoxal do ponto de vista bayesiano, é que, na verdade, o efeito de *e* com relação ao conteúdo indutivo de *h* seria sempre negativo. Com isso, Popper e Miller acreditam estar demonstrada a impossibilidade de qualquer indução probabilística e que toda inferência deve necessariamente ser dedutiva.

Howson e Urbach atribuem o resultado atingido por Popper e Miller ao fato de que estes empregaram uma definição nada ortodoxa para o excesso de conteúdo de uma sentença com relação a outra. Ao se tomar o conteúdo de uma hipótese *h* como o conjunto $Cn(h)$ de suas consequências não tautológicas, o excesso de conteúdo de *h* com relação a *e* pode ser definido como a diferença entre essas consequências e as consequências de *e*, isto é, $Cn(h) - Cn(e)$. Ao fazer isso, está-se utilizando a definição clássica para definir o excesso de A com relação a B, em que A e B são dois conjuntos: o excesso de A com relação a B corresponde ao maior subconjunto de A que não contém nada de B, isto é, ao resultado da diferença A − B.

No caso de Popper e Miller, os autores definem o excesso de *h* com relação a *e* como o maior subconjunto de $Cn(h)$ que não contém nada de $Cn(e)$ e que, adicionalmente, é também dedutivamente fechado. Com isso, o excesso de *h* com relação a *e* passa automaticamente a ser igual a $Cn(\sim e \vee h)$.



A razão para Popper e Miller estipularem que a diferença entre $Cn(h)$ e $Cn(e)$ seja um conjunto dedutivamente fechado é que eles optaram por determinar tal diferença no contexto algébrico do cálculo de sistemas dedutivos de Alfred Tarski (Howson e Urbach, 1993, p. 396). De fato, no contexto estipulado por Tarksi, as operações de soma, diferença, multiplicação e divisão operadas em conjuntos dedutivamente fechados devem, elas mesmas, ter como resultado conjuntos que sejam dedutivamente fechados. Nessas condições, a diferença entre *Cn(h)* e *Cn(e)* seria *Cn(*$\sim e \vee h$*).*

Howson e Urbach indicam que Popper e Miller não dão qualquer razão para definir a diferença entre *Cn(h)* e *Cn(e)* da maneira como o fazem, em lugar de empregar a definição com base na teoria dos conjuntos apresentada acima. Popper e Miller limitam a diferença entre *Cn(h)* e *Cn(e)* apenas ao conjunto das consequências de *h* que são também consequências de ~ *e*. Com isso, os autores deixam de lado todo um conjunto de consequências de *h*, aquelas que não são consequências de *e* nem de ~ *e*, como se vê com clareza na figura acima. Por essa razão, rejeitam a crítica de Popper e Miller, uma vez que seus pressupostos são arbitrários e sem justificativa adequada do ponto de vista bayesiano.

O PARADOXO DE MILLER

Na realização de inferências, o enfoque bayesiano faz uso de um princípio chamado de Princípio Principal (*Principal Principle*), segundo o qual a probabilidade subjetiva de um evento descrito pela sentença a será igual a r se, de acordo com as informações de fundo disponíveis, r for a probabilidade física (frequência) de a. De acordo com esse princípio, ressalvadas eventuais modificações nas condições de ocorrência do evento, a probabilidade subjetiva será equivalente à probabilidade física, isto é:

$$P[a/P^{*}\ (a)\ = r]\ = r$$

Na equação acima, P representa o grau de crença em determinado evento (função de probabilidade subjetiva) e P^{*} representa a probabilidade física (função de probabilidade física). Em seu argumento, Miller estipula, inicialmente, que *r* seja igual a $\frac{1}{2}$. Nessas condições, teremos que $P\left[a/P^{*}\ (a)\ =\ \frac{1}{2}\right]\ =\ \frac{1}{2}$, o que somente pode ocorrer se, e somente se, $P^{*}\,(a) =\ P^{*}\,(\sim a)$. O cálculo de probabilidade autoriza a substituição de sentenças equivalentes. Nessas condições, $P[a/P^{*}\ (a)\ =\ P^{*}\,(\sim a)]\ =\ \frac{1}{2}$. É possível, então, reescrever a primeira equação acima de tal forma que:

$$P[a/P^{*}\ (a)\ =\ P^{*}(\sim a)\,]\ =\ P^{*}(\sim a)$$

Para chegar a essa último resultado, não foi empregada qualquer consideração empírica. Embora o resultado não seja uma contradição, sua aplicação leva a resultados inconsistentes. É possível, por exemplo, efetuar as substituições acima a partir de $P^* (a) = \frac{2}{3}$, o que levará a que $P^* (a) = 2 P^*(\sim a)$, que contradiz o resultado obtido anteriormente.

Ao apresentar as deduções acima, Miller pretende mostrar que o chamado Princípio Principal é inválido e, com ele, todo o edifício bayesiano. Howson reconhece que as deduções de Miller, se verdadeiras, teriam um efeito desastroso para o bayesianismo. O cálculo da probabilidade posterior em hipóteses estatísticas depende do comportamento da função de verossimilhança $g(i) = P(e/h_i)$, em que h_i corresponde à família de hipóteses alternativas acerca da distribuição de determinada probabilidade física. Se o chamado princípio principal for inconsistente, será impossível utilizar tal função para se chegar às probabilidades posteriores. Howson e Urbach lembram que o edifício bayesiano está erigido sobre a noção de probabilidade subjetiva, que exige o emprego do Princípio Principal na atribuição de probabilidades. O conceito de probabilidade subjetiva, conforme desenvolvido por De Finnetti (e adotado por Howson e Urbach), fundamenta-se na noção de probabilidades como valor justo em que o indivíduo estaria disposto a apostar (*fair odds*).

Intuitivamente, é possível perceber que há algo errado com o raciocínio de Miller, mas o que, exatamente? Howson e Urbach (1993, p. 399) notam que o erro de Miller é difícil de identificar em função da notação empregada, que não explicita determinadas distinções teóricas. O equívoco está em substituir a segunda equação acima pela primeira, como se fossem equivalentes, uma vez que as duas equações fazem assertivas muito diferentes uma da outra. A primeira equação, isto é, $P[a/P^* (a) = r] = r$, é uma equação que envolve variáveis aleatórias (variáveis quantitativas cujos valores dependem de fatores aleatórios), como, por exemplo, o número de ocorrências da face *cara* em lançamentos sucessivos de uma moeda. Esse fato é obscurecido pela tendência a considerar $P^* (a)$ como um número e, com base nisso, efetuar as substituições. Na realidade, no contexto que se está analisando, $P^* (a) = r$ é uma sentença à qual a função P atribui determinado valor de probabilidade. $P^* (a)$ não é um número, mas uma expressão que pode assumir qualquer valor dentro do intervalo [0,1]. Quando determinada quantidade assume valores a depender dos estados possíveis do mundo, ela é uma função. Se os valores passíveis de serem assumidos estiverem associados a uma distribuição de probabilidades, trata-se de uma variável aleatória. Não é possível, assim, tomar um valor específico assumido pela função em determinado ponto e generalizar as conclusões extraídas, como faz Miller quando $r = \frac{1}{2}$ ou $r = \frac{2}{3}$. Com isso, Howson e Urbach acreditam ter deixado claro que as alegadas inconsistências do Princípio Principal não se sustentam (o princípio é válido), o que por sua vez mostra, uma vez mais, que o edifício bayesiano é consistente.

O PARADOXO DO INDÍCIO IDEAL

Tome-se, uma vez mais, uma moeda a ser lançada 1000 vezes. Um observador é requisitado a expressar sua probabilidade subjetiva quanto ao resultado *cara* do milésimo lançamento em duas ocasiões: inicialmente, sem que qualquer lançamento tenha sido feito; e após ter observado e registrado os 999 lançamentos anteriores. Sem qualquer informação de fundo, o referido observador, muito provavelmente, dirá que a probabilidade de o milésimo lançamento resultar em *cara* será $\frac{1}{2}$. Suponha-se, então, que após 999 lançamentos tenham sido registrados 503 resultados *cara*, o que é um indício de que a moeda em questão não está viciada. Nessas condições, o observador pode utilizar o Teorema de Bayes e atualizar a probabilidade prévia que, inicialmente, havia atribuído ao milésimo lançamento resultar em *cara*. Ocorre que, ao atualizar a probabilidade prévia, o valor da probabilidade posterior será praticamente o mesmo à luz dos resultados registrados, isto é, $\frac{1}{2}$. Ao condicionalizar o grau de crença, isto é, a probabilidade subjetiva com relação aos 999 lançamentos anteriores, a probabilidade posterior não foi afetada, a despeito da grande quantidade de indícios recolhidos.

Diante da situação descrita acima, Popper (1959, p. 407) criticou o enfoque bayesiano, o qual, ao expressar graus de crença por meio de um número, não se mostra capaz de distinguir, em determinadas situações, entre a probabilidade posterior após indícios terem se acumulado e a probabilidade prévia, que não toma qualquer indício em conta. A situação descrita acima, que constitui para Popper o chamado “paradoxo do indício ideal”, seria negativa para o enfoque bayesiano, pois demonstraria a necessidade de que os graus de crença em determinada proposição fossem representados em duas dimensões: a probabilidade da hipótese h à luz dos indícios disponíveis e a ponderação do conjunto de indícios em que a referida probabilidade está baseada.

Na situação em exame, a probabilidade prévia inicial (antes que qualquer lançamento tivesse sido realizado) estava baseada em considerações imprecisas. Por essa razão, o valor atribuível à ponderação do conjunto de indícios deveria ser baixo. Após 999 lançamentos, a probabilidade atribuída pelo observador ao resultado *cara* para o milésimo lançamento estava fundamentada em um conjunto considerável de indícios. Dessa forma, o valor desse conjunto de indícios deveria ser mais alto e, consequentemente, a probabilidade atribuída para o milésimo lançamento, embora coincida em valor com a probabilidade prévia, é mais confiável do ponto de vista epistêmico.

A resposta de Howson e Urbach a mais essa crítica de Popper é que o enfoque bayesiano é plenamente capaz de fazer o gênero de discriminação requerido, sem qualquer necessidade de desenvolver representações em duas dimensões para os graus de crença. Do ponto de vista bayesiano, o fato de a probabilidade prévia de $\frac{1}{2}$ ser estimada sem que qualquer lançamento tenha sido feito se reflete no fato de que a densidade de probabilidade é bastante difusa. A probabilidade de o lançamento resultar em *cara* constitui uma variável aleatória X. Antes de qualquer lançamento, o valor de X poderá variar em toda a amplitude do intervalo [0,1]. Como não foi feito qualquer lançamento, X pode assumir qualquer valor no intervalo e, consequentemente, a distribuição de probabilidades (os valores possíveis de serem assumidos por X) é difusa.

Após 999 lançamentos, a variável X estará concentrada de forma incisiva ao redor do valor $\frac{1}{2}$. Dessa forma, a diferença entre os valores ponderados, como quer Popper, dos indícios no primeiro e no segundo caso são expressos pela diferença na distribuição de probabilidades nas duas situações: difusa na primeira; concentrada na segunda. Do ponto de vista epistêmico, o enfoque bayesiano explica o fenômeno da seguinte forma: à medida que os lançamentos ocorrem, a frequência relativa de resultados *cara* irá convergir para um determinado valor, que é o mesmo valor para o qual se dará a convergência das probabilidades subjetivas condicionadas com relação aos resultados observados. Dessa maneira, os valores para a probabilidade subjetiva, ainda que permaneçam iguais do ponto de vista numérico, variam cada vez menos e são cada vez menos propensos a variar, a despeito dos novos indícios que se acumulam. Os valores se tornam, portanto, cada vez mais confiáveis do ponto de vista epistêmico e o enfoque bayesiano não tem qualquer dificuldade em explicar o fenômeno.

INDÍCIOS NÃO CONFIRMAM HIPÓTESES CONSTRUÍDAS PARA EXPLICÁ-LOS

Ao longo das páginas precedentes, foi mencionado que hipóteses, para serem aceitáveis e não serem descartadas como *ad hoc*, necessitam ser confirmadas por indícios independentes daqueles que elas foram construídas para explicar. Com base nessa condição, Ronald Giere retoma um argumento de Pierce, segundo o qual indícios não confirmam hipóteses construídas para explicá-los. É útil reproduzir a linguagem do próprio Giere sobre a questão:

> se os fatos conhecidos forem utilizados na construção do modelo e, com isso, tiverem sido embutidos na hipótese resultante... então a correspondência entre esses fatos e a hipótese não fornece qualquer indício de que a hipótese seja verdadeira, (uma vez que) esses fatos não tiveram qualquer oportunidade de refutar a hipótese[22].

O argumento, se verdadeiro, constituiria um golpe no edifício bayesiano. Conforme se viu no capítulo 1, porém, o fato de determinado indício ser utilizado para construir uma hipótese não significa, para o bayesiano, que seu impacto com relação à probabilidade prévia de *h* seja nulo. Nessa situação, espera-se que o impacto do indício *e* sobre a hipótese *h*, em lugar de ser nulo, seja apenas incremental, de modo a tornar a probabilidade posterior ligeiramente maior do que a probabilidade prévia.

Howson e Urbach ressaltam (Howson & Urbach, 1996, p. 408) que o argumento de Giere é inconsistente. Em primeiro lugar, há que se ter em conta que qualquer indício *e* confirma ou infirma a hipótese *h*, o que ocorre independentemente de *h* ter sido ou não construída a partir daquele indício específico. Howson e Urbach assinalam que o que Giere e outros defensores do argumento por ele proposto estão a fazer é confundir o aparato instrumental *E* com um de seus possíveis resultados, isto é, *e*. Uma vez que tal confusão é desfeita, fica fácil compreender porque o argumento de Giere não se sustenta: mesmo que *h* tenha sido construída a partir de um resultado específico de *E* (no caso, *e*), nada impede que o aparato instrumental *E*, caso tenha sido concebido de forma adequada, venha a produzir novos indícios, os quais infirmam *h*. Giere acredita que *h* sempre será confirmada, o que não é verdade, pois o aparato *E* poderá levar a resultados que infirmem aquela hipótese, independentemente de ela ter sido concebida a partir do resultado específico *e*.

---

[22] "if the known facts were used in constructing the model and were thus built into the resulting hypothesis… then the fit between these facts and the hypothesis provides no evidence that the hypothesis is true [since] these facts had no chance of refuting the hypothesis". (Giere, 1984, p. 161).

De acordo com o enfoque bayesiano, para que e confirme h, é necessário que o valor de $P(h/e)$ seja maior do que o valor de $P(h/{\sim}e)$. Não há qualquer razão para que isso não ocorra, mesmo que h tenha sido construída a partir de e. O fato de h ter sido construída tendo em vista e não significa, em absoluto, que e não confirma h, apenas que o efeito de confirmação ocorre de forma incremental.

O argumento de Giere possui uma versão um pouco mais sofisticada, mas que também se mostra passível de ser refutada (Howson & Urbach, 1993, p. 410). Suponha-se, novamente, que *h* foi construída com o auxílio do indício *e*. Suponha-se, ainda, que é possível decompor *e* em duas partes: *e´* (a parte de fato utilizada para compor *h*) e *e´´* (a parte restante do indício original). O argumento de Giere pode ser reelaborado de tal modo que se considerará que, no caso de *e* confirmar *h*, isto se dá apenas em função de *e´´* , que constitui a parte que não foi utilizada na construção da hipótese. Os defensores desse argumento chamam *e´´* de excesso explicativo (*explanatory surplus*) de *e* com relação a *h*. Como não há qualquer garantia de que *e´´* confirma *h*, pois não foi empregado na construção desta, qualquer efeito confirmatório de *e* com relação a *h* será devido precisamente a *e´´*. Assim, segundo essa versão aperfeiçoada do argumento de Giere, a única chance de um indício confirmar uma hipótese *h* é que tal indício possa refutá-la por meio de seu excesso explicativo, o que não ocorre se tiver sido empregado em sua construção.

Apesar de aparentemente válido, o argumento precedente também se mostra problemático. Isso ocorre porque se estão confundindo duas hipóteses distintas: a hipótese original *antes* de ter sido ajustada com base em *e´* (isto é, *h*) e a hipótese após os ajustes terem sido feitos (isto é, *h´*). É *h´*, não *h*, que é construída com base em *e´*. Com isso, o aparato instrumental *E* poderia ter produzido indícios que viessem a contradizer *h´*. O fato de *e´* confirmar *h´* não se mostra, portanto, problemático (Howson e Urbach, 1993, pp 410-11).

Não é verdade que uma hipótese *h* construída para acomodar determinado indício *e* nunca é por este confirmada. Tampouco é verdadeiro que *h* não é confirmada apenas por aquela parte de *e* que foi empregada em sua construção. Por outro lado, conforme se afirmou no início desta seção, sabe-se que o efeito de confirmação sobre *h* será maior caso se trate de indícios independentes, que não foram de alguma maneira "acomodados" pela hipótese. Nenhuma dessas constatações é alheia ao enfoque bayesiano, que se mostra capaz de oferecer explicações adequadas para cada uma delas.

Sem prejuízo das explicações precedentes, Howson e Urbach (1993, p. 411-412) oferecem uma demonstração mais geral das virtudes do bayesianismo com relação a essas críticas. Considere-se uma hipótese *h* em que alguns ou todos os seus parâmetros foram deixados indeterminados. Suponha-se que existe o indício *e* que é previsto pela hipótese rival $h´´$, mas somente pode ser previsto por *h* após seus parâmetros terem sido fixados com o auxílio de *e*. Uma vez fixados os parâmetros de *h* com base em *e*, estar-se-ia diante de uma versão alternativa da hipótese original, $h´$. Suponha-se, enfim, que *h* e $h´´$ têm probabilidades prévias iguais e que *e* é independente de *h*. Nessas condições, é fácil mostrar que a versão ajustada $h´$ será muito menos confirmada por *e*, que foi empregada em sua construção, do que $h´´$.

Uma vez que $h´´$ prediz a ocorrência de *e*, é possível mostrar que a diferença entre as probabilidades posterior e prévia dessa hipótese será dada por (Howson e Urbach, 1993, p.412):

$$P(h´´/e) - P(h´´) = \frac{P(h´´)[1 - P(e)]}{P(e)}$$

Da mesma forma, é possível mostrar que a diferença entre a probabilidade posterior e prévia de $h´$ é dada por:

$$P(h´/e) - P(h´) = \frac{P(h´)[1 - P(e)]}{P(e)}$$

Se a probabilidade de *e* for menor do que 1, o que é razoável de acordo com os axiomas da probabilidade que regem a aplicação do Teorema de Bayes, teremos que $P(h´) < P(h)$. Isso acontece uma vez que, como $h´$ constitui a versão aperfeiçoada de *h*, tendo em vista o indício *e*, então $P(h´) = P(h\&e) = P(h)P(e)$. Ao se multiplicar *P(h)* por um número menor que 1, tem-se um valor menor do que *P(h)*, o que torna evidente que $P(h´) < P(h)$. Ao se transpor esse resultado às equações acima, observa-se que, como, por hipótese, $P(h) = P(h´´)$, então

$$\frac{P(h´´)[1 - P(e)]}{P(e)} = \frac{P(h)[1 - P(e)]}{P(e)} > \frac{P(h´)[1 - P(e)]}{P(e)}$$

Com isso, temos que

$$P(h´´/e) - P(h´´) > P(h´/e) - P(h´)$$

O resultado mostra o que intuitivamente já se esperava, isto é, que o impacto de confirmação de *e* é muito maior com relação a $h´´$, do que com relação a $h´$. Não obstante, *e* confirma *h´*, ainda que em grau bem menor do que confirma *h´´*. O enfoque bayesiano explica de forma adequada, portanto, a questão. Um indício empregado na construção de uma teoria é capaz de confirma-la tanto quanto um indício chamado independente, apenas em grau menor.

A TEORIA DE DEMPSTER-SHAFER

O modelo bayesiano constitui, como se procurou mostrar até aqui, uma teoria matemática para o raciocínio inferencial. O bayesianismo não é, porém, o único modelo que procurou fornecer bases mais sólidas para o raciocínio indutivo. Nas últimas décadas, com os desenvolvimentos no campo da inteligência artificial e da computação, outros enfoques foram desenvolvidos. Uma dessas alternativas mostrou-se especialmente influente. Trata-se da chamada Teoria de Dempster-Shafer, a qual se fundamenta na noção de funções de crença (*belief functions*), que foi desenvolvida por Glenn Shafer (1976) em contraposição a alegadas insuficiências no enfoque subjetivista bayesiano, muito embora ela também parta de uma abordagem essencialmente subjetivista.

A Teoria de Dempster-Shafer constitui uma generalização do enfoque bayesiano, mas sem a utilização direta de probabilidades subjetivas. Enquanto o bayesianismo exige a atribuição de probabilidades para cada uma das questões relevantes em exame, as funções de crença na Teoria de Dempster-Shafer permitem a atribuição de graus de crença com relação a determinada sentença com base na probabilidade atribuída a questões relacionadas a ela. Os graus de crença, na formulação de Dempster-Shafer, não necessariamente apresentam as mesmas propriedades que caracterizam as probabilidades. Na verdade, o quanto os graus de crença irão se distinguir das probabilidades irá depender do quão próximas ou relacionadas são as duas sentenças em exame. A teoria está baseada, assim, em duas ideias principais: a determinação de graus de crença a partir das probabilidades subjetivas de uma situação epistêmica correlata à questão que se pretende examinar; e a combinação desses graus de crença baseados em indícios independentes.

Para ilustrar a primeira ideia, tome-se o exemplo, dado por Shafer, de uma pessoa que tem probabilidades subjetivas para o quão confiável é sua amiga Betty. A probabilidade de que Betty é confiável é 0,9; a probabilidade de que ela não é confiável é 0.1. Imagine-se, agora, que ela afirma que uma árvore caiu sobre meu carro. A assertiva deverá ser verdadeira se Betty for confiável. Ocorre, também, que a assertiva não será necessariamente falsa se Betty não for confiável (ela pode não ser confiável e, ainda assim, a assertiva ser verdadeira). O testemunho de Betty justifica o meu grau de crença estimado em 0,9 de que uma árvore, de fato, caiu sobre meu carro. O testemunho também justifica um grau de crença estimado em zero – e não estimado em 0,1– de que nenhuma árvore não caiu sobre meu carro. O grau de crença estimado em zero não significa que eu estou certo de que nenhuma árvore caiu sobre meu carro, como uma probabilidade de zero implicaria. Significa, tão somente, que o testemunho de Betty não me dá motivo para acreditar que nenhuma árvore caiu sobre meu carro. Nesse caso, a função de crença, que varia conforme as probabilidades atribuídas à confiabilidade de Betty, é constituída pelos valores 0,9 e zero, o que constitui situação distinta de caso estivesse baseada nos valores de probabilidade (0,9 e 0,1). O grau de confiabilidade de Betty constitui a questão epistêmica correlata cujas probabilidades são utilizadas para investigar a questão que se pretende examinar, que diz respeito à probabilidade de que uma árvore tenha caído sobre meu carro.

Para ilustrar a forma como a teoria combina diferentes graus de crença, imagine-se, agora, a situação em que eu atribuo uma probabilidade subjetiva de 0,9 com relação à confiabilidade de Sally, a qual também testemunha, independentemente de Betty, que efetivamente uma árvore caiu sobre meu carro. O evento "Betty é confiável" é independente do evento "Sally é confiável" e, nesse caso, a probabilidade de que ambas sejam confiáveis será 0,9 × 0,9 = 0,81. Da mesma forma, a probabilidade de que nenhuma das duas é confiável equivale a 0,1 × 0,1 = 0,01. A probabilidade de que ao menos uma das duas seja confiável é $1 - 0{,}01$, ou seja, 0,99. Uma vez que ambas afirmaram que uma árvore caiu em meu carro, o fato de ao menos uma delas ser confiável faz que eu atribua a esse evento o grau de crença de 0,99.

Suponha-se, agora, que Betty e Sally se contradigam – Betty afirma que uma árvore caiu em meu carro; Sally afirma que nenhuma árvore caiu em meu carro. Nesse caso, ambas não podem estar corretas e não podem ser as duas confiáveis. Apenas uma é confiável ou nenhuma é confiável. A probabilidade prévia de que apenas Betty é confiável equivale a 0,09, que constitui o resultado da multiplicação da probabilidade de Betty ser confiável (0,9) pela probabilidade de Sally não ser confiável (0,1). Da mesma forma, a probabilidade de que apenas Sally é confiável é igual a 0,09, resultante da multiplicação da probabilidade de Sally ser confiável (0,9) pela probabilidade de Betty não ser confiável (0,1). Por fim, a probabilidade de que nenhuma das duas seja confiável será 0,1 × 0,1 = 0,01. As probabilidades posteriores, dado que Sally e Betty não são ambas confiáveis, serão, respectivamente, $\frac{9}{19}$; $\frac{9}{19}$ e $\frac{1}{19}$. Com isso, tem-se um grau de crença equivalente a $\frac{9}{19}$ de que uma árvore caiu em meu carro porque Betty é confiável e um grau de crença de $\frac{9}{19}$ na mesma afirmativa porque Sally é confiável.

Em suma, a aplicação da teoria mostrou ser possível obter graus de crença para uma questão (uma árvore caiu sobre meu carro?) a partir das probabilidades atribuíveis a outras questões (é a testemunha confiável?). A implementação da teoria de Dempster-Shafer envolve dois elementos intimamente relacionados. Inicialmente, é necessário que a questão em exame seja articulada por meio de indícios independentes para que, em seguida, sejam efetuados os cálculos exigidos pela teoria. Se Betty e Sally testemunharem, de forma independente, que ouviram um ladrão entrar em minha casa, é possível que ambas tenham confundido o ruído produzido por um cachorro com o ladrão. Em função dessa incerteza compartilhada entre Betty e Sally, não é possível combinar graus de crença baseados nos indícios de cada uma conforme exige a regra de Dempster. Por outro lado, se for considerada explicitamente a possibilidade de haver um cachorro, então torna-se possível identificar três indícios independentes: o indício a favor ou contra a presença do cachorro, o indício relativo à confiabilidade de Betty e o indício relativo à confiabilidade de Sally. Com isso, torna-se possível combinar esses indícios conforme a regra de Dempster e os cálculos correspondentes são facilitados pela forma como a questão em exame foi articulada, isto é, a partir de três indícios independentes.

Qual a relação entre a teoria de Dempster-Shafer e o bayesianismo? A combinação de indícios distintos como meio para determinar o grau de crença em uma sentença é apresentado como uma alternativa mais adequada do que o bayesianismo para explicar a confirmação ou infirmação de determinada hipótese. O problema é, conforme ressaltam Howson e Urbach (1993, pp. 426-427), que inexiste qualquer garantia de que os indícios e respectivas crenças por estes evocados irão satisfazer os condicionamentos impostos por Shaffer. É como, afirmam Howson e Urbach, se a teoria proposta por eles apresentasse uma sintaxe mas fosse, ainda, desprovida de semântica. O contraste com o bayesianismo não poderia ser mais intenso, na visão dos autores, uma vez que naquele enfoque, a confirmação de uma hipótese é explicitamente medida como modificação em graus de crença induzida por indícios especificados de forma clara. Além disso, o bayesianismo repousa sobre a premissa de que os graus de crença, se forem consistentes, apresentam a estrutura formal de probabilidades. A teoria de Dempster-Shafer define confirmação de forma distinta. Howson e Urbach insistem, não obstante, que no bayesianismo é possível definir uma estrutura formal e justificá-la de forma não ambígua.

Shafer (1981) chegou a propor uma justificação mais detalhada para sua teoria com base no que ele chamou de "exemplos canônicos" (*canonical examples*). Nesses exemplos, os indícios são considerados como mensagens codificadas ao acaso. É difícil compreender, porém, os motivos por trás do artifício de interpretar indícios como mensagens codificadas ao acaso, bem como de que maneira elas poderiam justificar a sintaxe proposta por Shafer. Ao propor que os indícios sejam considerados como mensagens codificadas ao acaso, Shafer estava apresentando uma alternativa ao que considerava uma falha crucial do bayesianismo, que interpreta os fenômenos que se propõe a explicar como situações derivadas de um mecanismo probabilístico ao acaso. São esses mecanismos que permitem chegar às distribuições de probabilidades com relação aos indícios disponíveis. Howson e Urbach reconhecem que qualquer distribuição de probabilidades pode ser modelada como um mecanismo probabilístico, por exemplo, de retiradas ao acaso de uma urna. Isso não significa que qualquer fenômeno deva, necessariamente, ser compreendido dessa forma, nem essa seria uma exigência do enfoque bayesiano. Para o enfoque bayesiano, as distribuições de probabilidade nada mais são do que distribuições de graus de crença.

Shafer critica o bayesianismo com base no argumento de que as distribuições de probabilidades prévias não são baseadas em indícios específicos. Howson e Urbach reconhecem que, de fato, as distribuições prévias de probabilidades não estão baseadas na condicionalização obtida a partir de um indício inicial (Howson & Urbach, 1993, pp. 423-430). Argumentam, porém, que o bayesianismo nunca se propôs a justificar a distribuição prévia de probabilidades em bases epistêmicas estritas. Tal crítica foge ao propósito do enfoque bayesiano, que é simplesmente o de fornecer uma teoria para a realização de inferências a partir das probabilidades prévias para as probabilidades posteriores. As probabilidades prévias, para o bayesianismo, são elementos dados que servem de ponto de partida para se chegar a probabilidades posteriores, tendo em vista o impacto confirmatório ou infirmatório de determinado indício. Criticar o bayesianismo como o faz Shafer é repreender o enfoque por algo que, conscientemente, é deixado de fora.

A crítica mais intensa de Shafer é de que o bayesianismo, ao estar baseado no cálculo de probabilidades, não é capaz de representar graus de incerteza de forma adequada. A objeção é de que nenhuma teoria baseada no cálculo de probabilidades pode representar adequadamente um estado de ignorância entre alternativas. Com efeito, os próprios Howson e Urbach indicaram suas reservas ao princípio da indiferença, conforme apresentado no capítulo 1. Shafer acredita que a forma adequada de representar um estado de ignorância total é atribuir grau de crença zero a cada uma das alternativas, o que é possível de acordo com a teoria que propõe. No cálculo das probabilidades, por outro lado, o estado de ignorância completo pode ser expresso por uma distribuição uniforme e diferente de zero para cada alternativa em questão.

Embora reconheçam que há mérito na crítica de Shafer, Howson e Urbach acreditam que ela não é, nem de longe, suficiente para desacreditar o enfoque bayesiano. Isso porque a própria compreensão do que é um estado de ignorância não é totalmente clara. O cálculo de probabilidades é perfeitamente compatível com uma situação em que o estado de ignorância não é idêntico com relação às várias alternativas em questão e isso é, insistem Howson e Urbach, intuitivamente correto. Se, diante de um exercício de retirada de dez bolas numeradas, um observador atribui a mesma probabilidade $(\frac{1}{10})$ a cada um dos resultados, então ele não terá o mesmo grau de crença com relação às assertivas "será retirada a bola zero" e "será retirada uma bola diferente de zero". Na teoria de Shafer, por outro lado, é possível atribuir a mesma probabilidade a ambas assertivas, o que constitui uma violação do cálculo de probabilidades e, como insistem Howson e Urbach, uma mostra de que aquela teoria não se sustenta. Da mesma forma, a teoria de Shafer-Dempster é compatível com a negação de que o aumento no grau de crença em h deva ser acompanhado pela diminuição no grau de crença de $\sim h$. Isso seria, insistem Howson e Urbach, mais uma mostra de que a teoria é inconsistente e carece de qualquer interpretação coerente.

CONCLUSÃO

Neste capítulo, procurou-se apresentar as críticas mais comuns e incisivas ao enfoque bayesiano para o raciocínio científico dentre as analisadas e respondidas por Howson e Urbach. A apresentação foi dividida em duas partes. Na primeira parte, foram examinadas as objeções que Howson e Urbach consideram as mais relevantes (subjetivismo e o problema do indício antigo) e que, por essa razão, fazem parte da última edição do livro *Scientific Reasoning – the Bayesian Approach.* Na segunda parte, foram analisadas as críticas de Popper e Giere, bem como as objeções levantadas pelo enfoque alternativo proposto por Glenn Shafer. Esse segundo conjunto faz parte das edições anteriores, de 1989 e 1993, de *Scientific Reasoning – the Bayesian Approach* e foram abandonadas na terceira edição do livro, uma vez que Howson e Urbach consideram, como deixam claro no prefácio à edição de 2006, que tais objeções não se sustentam e já haviam sido respondidas de maneira adequada.

Ao responder a todo o conjunto de objeções e ressalvas explicitado acima, Howson e Urbach insistem na consistência do raciocínio indutivo fundamentado no cálculo de probabilidades. Uma vez aceito o fato de que graus de incerteza podem ser expressos por meio de probabilidades, a obediência aos axiomas do cálculo de probabilidades se mostra como o caminho natural para o desenvolvimento de inferências fundamentadas nos indícios provenientes da pesquisa científica.

Iniciado no século XVII, o desenvolvimento do cálculo de probabilidades floresceu nos séculos seguintes e se tornou a base para desenvolvimentos em campos tão diversos quanto as teorias sobre ação racional, processo decisório e raciocínio indutivo. Problemas e paradoxos foram, porém, identificados – especialmente os relacionados com o Princípio da Indiferença –, o que levou autores como Fisher e Popper a defenderem o abandono do programa de pesquisa como um todo.

Não foi senão com os trabalhos de Ramsay, von Neumann, Morgesntern e De Finetti que uma resposta adequada pôde ser desenvolvida e que, na avaliação de Howson e Urbach, logrou apresentar uma teoria para as probabilidades epistêmicas que não estivesse limitada por princípios pretensamente objetivos – caso do Princípio da Indiferença –, e impusesse parâmetros consistentes para a realização de inferências. Sem prejuízo de que as probabilidades sejam subjetivas, os parâmetros a elas impostos na realização de inferências é perfeitamente objetivo. O propósito de Howson e Urbach, ao escrever *Scientific Reasoning*, era precisamente o de convencer os adeptos da teoria objetiva das probabilidades de que não há nada de subjetivo no enfoque bayesiano *como teoria da inferência* (Howson e Urbach, 1993, p 438).

No próximo capítulo, serão discutidas outras críticas ao bayesianismo, as quais não foram objeto de resposta direta por Howson e Urbach. Com isso, os limites de aplicação do bayesianismo ficarão mais claros e se poderão compreender melhor as razões que levaram determinados filósofos a buscar enfoques alternativos e, alegadamente, mais objetivos. Sem prejuízo das dificuldades práticas de aplicação do enfoque bayesiano a determinadas situações, o grau de subjetivismo advindo da interpretação subjetivista das probabilidades sempre constituiu um elemento de intranquilidade para um grande número de autores. É a rejeição da possibilidade de analisar o raciocínio científico em bases subjetivistas que está na base do falsificacionismo popperiano e, mas recentemente, da estatística do erro, conforme se verá.

3. AS CRÍTICAS NÃO CONSIDERADAS POR HOWSON E URBACH

Neste capítulo pretende-se dar continuidade ao capítulo anterior e apresentar críticas ao enfoque bayesiano que não foram objeto de resposta direta por Colin Howson e Peter Urbach em qualquer das edições de *Scientific Reasoning – the Bayesian Approach*, nem tampouco em publicações posteriores. A partir dessas críticas, será possível observar com maior clareza os limites e virtudes desse enfoque como instrumento para analisar o raciocínio científico.

O DEBATE BAYESIANISMO VERSUS FALSIFICACIONISMO SEGUNDO GILLIES

Em junho de 1990, Donald Gillies publicou o artigo *Bayesianism versus Falsificationism* (Gillies, 1990) em que, com o pretexto de apresentar uma resenha do livro de Howson e Urbach, retoma o clássico embate entre os enfoques bayesiano e falsificacionista. Gillies reconhece a impossibilidade de examinar em detalhe cada um dos argumentos desenvolvidos por Howson e Urbach, inclusive à luz do propósito dos dois autores de não apenas apresentar o panorama mais completo possível do enfoque por eles defendido, mas, também, de demonstrar a superioridade deste sobre o falsificacionismo. Gillies, que não esconde sua simpatia pelo falsificacionismo nesse debate, concentra-se em duas questões específicas: a) as hipóteses *ad hoc*; e b) o papel dos testes de significância no raciocínio científico.

Tome-se a seguinte situação descrita por Howson e Urbach no capítulo 4 de *Scientific Reasoning* (2006, pp. 106-115): combina-se uma teoria principal *t* com a hipótese auxiliar *a* (considere-se, por comodidade, *a* equivalente à conjunção de várias hipóteses auxiliares $a_1, a_2, a_3, \dots, a_n$), com o objetivo de predizer o indício *e*. Ao serem realizados os experimentos correspondentes, observou-se, porém, a ocorrência de *e´*, que é incompatível com *e*. De modo a manter intacto o poder explicativo da teoria principal *t*, uma nova hipótese auxiliar, *a´*, é apresentada, a qual, em conjunção com *t*, permite a almejada previsão do indício *e*. A questão central, conforme sublinha Gillies, é se, na situação descrita, *a´* deve ser considerada uma hipótese *ad hoc* ou não.

Do ponto de vista falsificacionista (Popper, 1972, capítulo IV, seção 20), a introdução de hipóteses auxiliares é aceitável se não acarretar a diminuição do grau de falsificabilidade do sistema em exame mas, ao contrário, se ampliá-lo. Segundo Popper, a introdução de nova hipótese auxiliar deve ter, como efeito, tornar a teoria mais falsificável, pois ela deve passar a proibir mais do que anteriormente. Uma hipótese auxiliar não é considerada *ad hoc* e é, portanto, admissível se sua introdução levar à identificação de novos indícios, os quais podem ser comparados com os resultados originalmente obtidos. Uma hipótese auxiliar é considerada *ad hoc* – portanto, inadmissível –, se sua introdução nada mais faz do que restabelecer a concordância entre a teoria principal e os indícios obtidos. Um exemplo bem sucedido da introdução de uma hipótese auxiliar foi, precisamente, a postulação da existência do planeta Netuno por Adams e Leverrier, que permitiu não apenas explicar de forma satisfatória as anomalias na órbita de Urano, como também levou a uma série de indícios independentes, que foram registrados pelos cientistas posteriormente.

Howson e Urbach defendem que o critério usual para distinguir hipóteses auxiliares admissíveis de hipóteses *ad hoc* é inadequado. Isso acontece, no entender daqueles autores, uma vez que tal critério poderia levar ao descarte de hipóteses auxiliares perfeitamente satisfatórias como se elas fossem *ad hoc.* Como em tantas ocasiões em *Scientific Reasoning*, Howson e Urbach defendem seu argumento por meio da construção de situações hipotéticas.

Tome-se, dizem eles, a hipótese *h* segundo a qual uma urna contém apenas bolas vermelhas. Um experimento é realizado em que bolas são retiradas, sua cor anotada e, a seguir, recolocadas na urna. O procedimento é repetido 10.000 vezes. O resultado é que 4.950 das bolas selecionadas são vermelhas; o restante das bolas apresenta a cor branca. A partir desse resultado experimental, é perfeitamente razoável substituir a hipótese original por outra, segundo a qual a urna em questão contém bolas vermelhas e brancas em quantidades aproximadamente iguais. A revisão da hipótese original *h* foi feita, entretanto, sem que qualquer indício independente tenha sido registrado: a hipótese foi modificada apenas com base nos mesmos indícios que desacreditaram a hipótese original.

Gillies considera falho o argumento de Howson e Urbach para desacreditar o critério de classificação de hipóteses auxiliares como *ad hoc.* De fato, a hipótese reformulada *h´*, segundo a qual a urna contém bolas vermelhas e brancas em quantidades similares, apresenta uma série de consequências que podem, uma vez articuladas em experimentos, originar indícios independentes. Seria possível, por exemplo, esvaziar a urna e contar a quantidade de bolas vermelhas e brancas para verificar que sua quantidade é aproximadamente igual. De forma semelhante, caso se repetisse o experimento das 10.000 retiradas de bolas da urna, seria possível, com base na hipótese reformulada $h´$, efetuar predições acuradas acerca da quantidade de bolas vermelhas e de bolas brancas. Esses indícios seriam independentes e os experimentos que os originam constituem consequência direta da reformulação da hipótese h.

Para Gillies, o exemplo hipotético de Howson e Urbach não se sustenta ainda por outra razão. Isso porque a hipótese h (a urna contém apenas bolas vermelhas) não é refutada por todo o conjunto de indícios $e´$ que resulta das 10.000 retiradas. Basta a retirada da primeira bola branca para que a hipótese h seja refutada. Podemos dizer que o conjunto $e´$ é composto por $e´_1$ (o resultado correspondente à retirada da primeira bola branca) e por $e´_2$ (o restante dos indícios). Nessas condições, prossegue Gillies, é razoável tomar $e´_2$ como uma forte indicação de que a hipótese reformulada $h´$ não é uma hipótese *ad hoc*: se $h´$ fosse *ad hoc*, ela apenas daria conta de $e´_1$. Ocorre, porém, que $h´$ tem como consequência todo o conjunto de indícios independentes agrupados como $e´_2$.

Após defender o critério falsificacionsita para a introdução de hipóteses auxiliares, Gillies volta-se, a seguir, contra a interpretação que o enfoque bayesiano propõe para as hipóteses auxiliares e hipóteses *ad hoc*. Com isso, o autor quer mostrar que o critério falsificacionista não apenas não é ameaçado pelas críticas de Howson e Urbach, como também constitui uma alternativa superior ao enfoque bayesiano nessa questão.

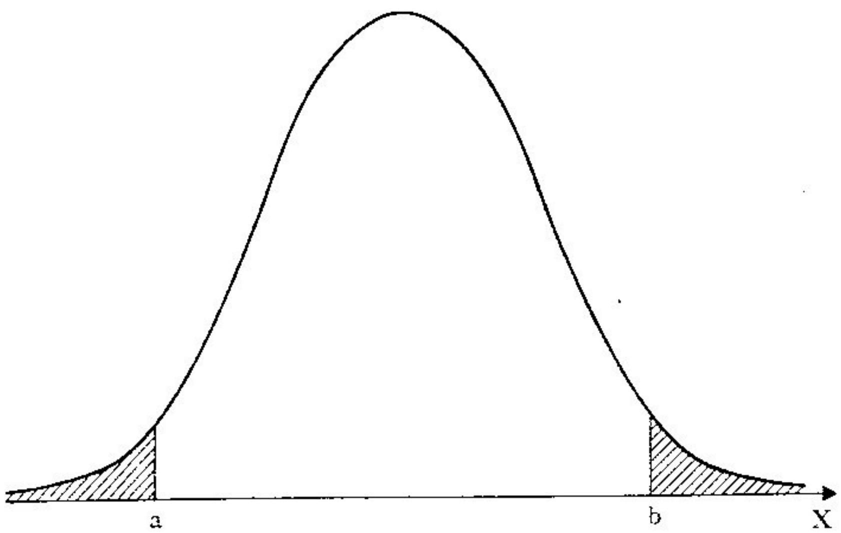
Para Howson e Urbach, uma hipótese auxiliar $a´$ será considerada *ad hoc* se $P(a´/e´\& b) \leq 0{,}5$ (Howson e Urbach, 1989, p.110). Nessa definição, $e´$ constitui o indício novo que refutou a hipótese predecessora de $a´$ e b expressa o conjunto de informações de fundo relevantes para a questão em exame. Gillies tenta mostrar que tal critério, baseado na condicionalização bayesiana, é de difícil aplicação.

Tome-se, uma vez mais, o exemplo aludido anteriormente acerca da introdução da hipótese auxiliar da existência de Netuno para explicar as anomalias na órbita de Urano. Viu-se que o critério de Popper explica de maneira satisfatória as razões pelas quais aquela hipótese é adequada e não deve ser descartada como *ad hoc*. Examine-se agora a maneira de o bayesianismo tratar a mesma questão. Para aplicar o critério de Howson e Urbach, é necessário estimar o valor de $P(a´/e´\& b)$ e, em seguida, verificar se tal valor é menor do que 0,5 (hipótese em que a hipótese deve ser descartada como *ad hoc*) ou maior do que 0,5 (hipótese em que não será considerada *ad hoc*). Como em tantas outras ocasiões de aplicação do enfoque bayesiano, a questão que se coloca é a necessidade de estimar a probabilidade de determinada expressão, no caso, de $a'/e' \& b$. Que parâmetros poderiam ser empregados para concluir se $P(a'/e' \& b)$ é maior ou menor do que 0,5? A resposta, para Gillies, é que simplesmente não há maneira satisfatória de estimar essa probabilidade. Qualquer tentativa de fazê-lo é desprovida de qualquer elemento objetivo. O cálculo em bases puramente subjetivas, por sua vez, tampouco representa qualquer garantia com relação à natureza da hipótese a'. Isso significa, ressalta o autor, que a aplicação do critério bayesiano não permite concluir de forma correta se a hipótese auxiliar é, ou não, *ad hoc.*

Além do problema da indecidibilidade, o critério de Howson e Urbach também sofre de um segundo vício, segundo Gillies: tal critério poderia levar à conclusão equivocada de que a introdução da hipótese de Netuno seria *ad hoc.* Isso ocorre porque a expressão *b*, que denota as informações de fundo relevantes para a questão em exame, certamente contém um amplo conjunto de indícios, nenhum dos quais indica a existência de um planeta além de Urano. A proposição de um novo planeta (no caso, Netuno) iria além de todo o conjunto de indícios disponíveis, o que certamente constituiria um argumento em favor de se tratar de uma hipótese *ad hoc.* Além disso, as perturbações na órbita de Urano podem ser explicadas por uma série de hipóteses alternativas à existência de outro planeta, como, por exemplo, o abandono das leis de Newton a grandes distâncias do sol ou a atuação de forças não-gravitacionais. Imagine-se que as anomalias na órbita de Urano poderiam ter sido explicadas por cinco diferentes hipóteses, entre as quais a existência de Netuno. De acordo com os axiomas do cálculo de probabilidades, é natural, nessa situação em que não há novos indícios disponíveis, que cada uma dessas hipóteses apresente probabilidade menor do que 0,5, para que a soma de todas elas seja igual a 1. Com isso, a hipótese auxiliar relativa a Netuno seria, de acordo com o critério proposto por Howson e Urbach, considerada *ad hoc*, o que definitivamente não é o caso.

Em suma, a defesa de Gillies do falsificacionismo como método para avaliar se uma hipótese é, ou não, *ad hoc,* quando comparada ao bayesianismo, baseia-se em dois argumentos complementares: a) o falsificacionismo não é afetado pela crítica que lhe fazem Howson e Urbach; e b) o falsificacionismo constitui uma alternativa melhor do que o bayesianismo, pois este é de difícil aplicação e pode, ainda, levar a resultados equivocados, como no caso da hipótese de Adams e Leverrier quanto à existência do planeta Netuno para explicar as anomalias na órbita de Urano.

Examine-se agora o segundo tópico escolhido por Gillies para defender o falsificacionismo frente à alternativa bayesiana: a aplicação dos testes de significância (*significance tests*). O chamado teste de significância nada mais é do que um procedimento para testar hipóteses. O procedimento exige que se deduza, de uma hipótese h, uma variável aleatória X, a qual possui uma distribuição aleatória de probabilidades. Dois pontos a e b são escolhidos de modo a dividir a figura que expressa a distribuição de probabilidades em uma parte central e duas caudas, de modo que $a \leq X \ \leq b$, ou seja, $X > a$ ou $X < b$. As caudas da curva são fixadas de modo que a probabilidade de se obter um resultado na região por elas demarcada, dada a veracidade de h, tem um valor baixo que é determinado, precisamente, pelo parâmetro pré-fixado do teste de significância. Normalmente, esse valor é fixado entre 1% e 10%. O mais comum é que seja fixado em 5%.



Um indício cujo valor esteja localizado na região delimitada pelas caudas é considerado uma instância de falsificação de h. Por meio da fixação dos parâmetros aludidos, o teste permite medir a probabilidade de se considerar h falsificada quando a hipótese é, na verdade, verdadeira. O escopo do teste é, assim, claramente falsificacionista. Quais seriam os argumentos de Howson e Urbach contra o teste de significância?

Howson e Urbach empregam dois argumentos independentes para desacreditar os testes de significância. O objetivo é mostrar que esses testes podem levar a resultados opostos, seja em função da forma como os indícios são combinados e tratados, seja em função da regra de parada (*stopping rule*) adotada para o experimento.

Suponha-se que determinado experimento repetível tem, como resultados, os números 0 ou 1. Suponha-se, ainda, que a hipótese h é de que as repetições são independentes e $P(0) = P(1) = \frac{1}{2}$. Após o experimento ter sido repetido um número $n = 2r$ de vezes, foram registrados os indícios $e_1 = 0, e_2 = 1, e_3 = 0, e_4 = 1, \ldots, e_{2i-1} = 0, e_{2i} = 1, \ldots, e_{2r-1} = 0, e_{2i} = 1$. Defina-se agora duas variáveis aleatórias X e Y, de modo que $X = \frac{e_1 + e_2 + \cdots + e_n}{n}$ e $Y = \frac{e_2 + e_4 + \cdots + e_{2i} + e_{2r}}{r}$. Nessas condições, $X = \frac{1}{2}$ e um teste de significância baseado em *X* irá corroborar *h*; $Y = 1$ e um teste de significância baseado em *Y* irá refutar *h*.

Gillies afirma que não há nada de paradoxal com a situação descrita, diferentemente do que pretendem Howson e Urbach. Isso porque X e Y extraem diferentes informações do conjunto de indícios recolhidos. Esse conjunto de indícios produz uma frequência de resultados iguais a "1" que é condizente com *h*, como mostra o teste conduzido a partir da variável *X*. Ao mesmo tempo, o observador imediatamente reconhece que *h* não pode ser verdadeira, pois a sequência entre resultados "0" e "1" é alternada e não aleatória, como seria de se esperar. A regularidade na alternância de resultados mostra que as repetições não são independentes umas das outras, contrariamente ao que se supunha. A variável *Y* expõe essa regularidade ao selecionar apenas os resultados pares, o que modifica a frequência de resultados "1" de $\frac{1}{2}$ para 1. Com isso, *Y* corretamente refuta *h*.

Ao apresentar sua defesa dos testes de significância, o que Gillies está implicitamente dizendo é que o cientista que os empregar não deverá realizar apenas o teste baseado na variável *X* ou o teste baseado na variável *Y*, e por meio do enfoque bayesiano, concluir pela confirmação ou infirmação da hipótese. Esse seria o procedimento para Howson e Urbach e constituiria um erro. O cientista utilizará os testes de significância inúmeras vezes, com base em variáveis distintas, com vistas a testar a hipótese de forma tão rigorosa quanto possível. Ao testar a hipótese de forma rigorosa, ele terá condições de examinar a discrepância e resultados e concluir, de forma informada, pela corroboração ou refutação de sua hipótese.

O segundo argumento de Howson e Urbach contra os testes de significância está baseado na adoção, por experimentos diversos, de distintas regras de parada. Imagine-se que a hipótese *h* é a de que uma moeda não é viciada e que seus lançamentos são independentes. Um primeiro experimento é realizado, o qual consiste em lançar a moeda vinte vezes e anotar os resultados. O segundo experimento para testar *h* é distinto e consiste em lançar a moeda tantas vezes quanto necessário até que sejam registrados seis resultados "cara". O experimento é abandonado ao se obter o sexto resultado "cara". Em geral, os experimentos darão resultados divergentes, mas nada impede que produzam exatamente o mesmo resultado. É possível, por exemplo, que sexto resultado "cara" no segundo experimento só seja obtido no vigésimo lançamento. Nesse caso, ter-se-ia um resultado de seis "caras" e catorze "coroas", cuja ordem poderia ser exatamente a mesma produzida pelo primeiro experimento.

Howson e Urbach efetuam uma série de cálculos (1989, pp.170-171) e mostram que o resultado seis "caras" e catorze "coroas" refuta *h* no segundo experimento no nível de significância de 5%; o mesmo resultado corrobora *h* no primeiro experimento no nível de significância também de 5%. A conclusão, para os autores, é que, no enfoque clássico, a inferência a ser extraída de um procedimento de teste de significância depende da regra de parada escolhida. Com isso, o teste de significância não seria um procedimento confiável em condições de fornecer informações úteis para o analista acerca da falsificabilidade de *h*. Como Gillies responde a isso?

A resposta de Gillies consiste em contrastar as visões falsificacionista e bayesiana do raciocínio científico. O resultado aludido é paradoxal para Howson e Urbach porque eles são bayesianos. Tal resultado não representa um problema para o falsificacionista. Para o bayesianismo, o objetivo do raciocínio científico é realizar inferências acerca da probabilidade de uma hipótese a partir de um conjunto de indícios cuja forma de obtenção é totalmente irrelevante. Um falsificacionista, por sua vez, formula uma hipótese *h* e procura testá-la com relação a um conjunto de indícios. Ele tenta refutar *h* por meio de experimentos tão rigorosos quanto possível. Caso *h* não seja refutada pelos testes laboratoriais, ela pode ser aceita provisoriamente. Para o falsificacionista, o método experimental é crucial e os indícios apenas têm significado no contexto específico do experimento em que foram produzidos. Assim, para o falsificacionista, é natural que um mesmo indício tenha significados distintos – e leve a conclusões distintas – a depender do experimento em que tenha sido registrado. Qualquer conclusão acerca da hipótese apenas será tirada após o cotejamento das conclusões parciais de cada experimento. E o cotejamento dos resultados dos vários experimentos deixará claro se a hipótese deve ser conservada ou descartada. O bayesiano, por sua vez, não está interessado no contexto experimental. Ele se concentra exclusivamente nos indícios e nas inferências que pode realizar a partir deles.

No exemplo específico mencionado, o significado do resultado seis "caras" exibe significados bem diferentes em cada situação. No segundo experimento (lançamento de uma moeda até ser registrado o resultado seis "caras"), o resultado não apresenta qualquer significado especial, pois sempre será obtido. Obter seis "caras" é, aliás, o objetivo do experimento. No primeiro experimento (lançamento da moeda vinte vezes), por outro lado, o resultado é bastante significativo, uma vez que não se sabe ao certo se será obtido ou não.

Após criticar as regras de parada como responsáveis pela produção de resultados supostamente inconsistentes nos testes de significância, Howson e Urbach (1989, p.171) também criticam tais regras por violar a objetividade que deve guiar a execução do empreendimento científico. Para eles, haveria um risco intrínseco nas regras de parada, que se consubstancia no fato de elas poderem ser utilizadas para "atender a certas intenções pessoais do pesquisador".

Ao responder a essa alegação, Gillies lembra que uma das principais características do método científico é que um experimento realizado pelo cientista A deve ser repetível pelo cientista B. Nenhuma "intenção pessoal" deve, portanto, estar envolvida na concepção de determinado experimento. No caso específico das regras de parada, elas devem ser fixadas de forma clara. De fato, essa última crítica de Howson e Urbach aos testes de significância parece destinada, antes, aos cientistas que se desviam do código de conduta do que aos testes propriamente ditos.

Em suma, Howson e Urbach efetuam uma descrição acurada dos testes de significância em seu intento de criticá-los. A questão é, porém, que as "falhas" por eles apontadas não constituem falhas para um falsificacionista, mas decorrências naturais da busca incessante pela falsificação de uma hipótese *h*. O falsificacionista, ao ser defrontado com uma situação em que a hipótese *h* passou determinado teste de significância baseado em um experimento, buscará conceber novos experimentos e testes para comprovar ainda com maior certeza se a hipótese deve, ou não, ser provisoriamente aceita. Caso isso resulte na refutação da hipótese, ele deverá procurar entender as razões por trás de tal resultado e somente então decidir se é possível aceitar *h* ou se novos experimentos devem ser conduzidos. Não há nada de equivocado nisso do ponto de vista falsificacionista, por mais que Howson e Urbach não estejam confortáveis com a ideia de que diferentes testes podem levar a conclusões diametralmente opostas quanto à corroboração ou refutação de h.

O BAYESIANISMO E A RIGIDEZ DO ENFOQUE TEÓRICO

Em artigo publicado em 2001, Donald Gillies efetuou novo exercício de comparação entre o bayesianismo e o falsificacionismo (2001, pp. 363-379). Nessa oportunidade, Gillies procurou identificar e expor os limites de aplicação do bayesianismo. Para ele, o bayesianismo somente pode ser aplicado de forma válida e coerente a situações em que há um quadro teórico (*theoretical framework*) conhecido e que se supõe não será alterado no curso da investigação. Em situações em que se verifica alteração no quadro de referência teórico, haveria modificações nos valores de probabilidades que não respeitariam as exigências do Teorema de Bayes e da condicionalização bayesiana. Frente a uma situação que não é bem conhecida pelo investigador, seria preferível utilizar procedimentos de investigação falsificacionistas, em lugar de recorrer ao bayesianismo. Vejamos em mais detalhes o argumento de Gillies.

O argumento de Gillies de que o bayesianismo somente funciona a contento em situações em que o quadro teórico é fixo desenvolve-se com base em dois exemplos: a investigação de Neyman acerca da distribuição de larvas em pequenos lotes de terra e as investigações de De Finetti acerca da permutabilidade entre variáveis aleatórias. Com base nesses exemplos, Gillies concluirá ser imprescindível ter certeza acerca do quadro teórico correspondente para que o enfoque bayesiano possa ser utilizado na interpretação de resultados experimentais.

A investigação de Neyman estava relacionada a estudos para o controle de pragas em plantações. Determinada plantação é dividida, para fins de controle, em pequenas parcelas. A seguir, conta-se a quantidade de larvas encontradas em cada parcela. Obviamente, a quantidade de larvas é variável conforme a parcela de terreno que se considere. Inicialmente, Neyman acreditava que a distribuição de probabilidades da localização das larvas seguia o modelo de Poisson. Segundo tal modelo, a probabilidade p_n de haver *n* larvas em uma determinada parcela de terreno é dada pela fórmula $p_n = \frac{e^{-\lambda} \times \lambda^n}{n!}$, em que e é a base do logaritmo natural (e = 2,71828...); *n!* é o fatorial de n ($n \times n-1 \times n-2 \times ... \times 2 \times 1$); e λ é um número real, igual ao número esperado de ocorrências em determinado espaço de tempo.

Neyman acreditava ser razoável que a distribuição de larvas nas parcelas de terreno seguisse o modelo de Poisson. Em um experimento anterior similar, ele havia empregado essa hipótese, a qual foi confirmada pelos indícios recolhidos. Neyman realizou experimentos e submeteu a hipótese h do modelo de Poisson a vários testes de significância. Surpreendentemente, os resultados dos testes de significância indicaram que a hipótese deveria ser descartada.

Até o momento em que sua hipótese foi refutada, Neyman não havia considerado qualquer alternativa com relação à probabilidade de distribuição das larvas. De sua experiência anterior, tudo parecia indicar que o modelo de Poisson se aplicava à situação em exame. Foi somente com a falsificação da hipótese que Neyman passou a contemplar hipóteses alternativas. Ele procurou conhecer melhor o comportamento das larvas, tendo consultado um especialista no tema. Com isso, chegou à conclusão de que as larvas deveriam estar distribuídas em grupos ao redor dos locais onde teriam sido depositados lotes de ovos. Os pontos em que os lotes de ovos teriam sido colocados seguiria o modelo de distribuição de Poisson, mas não a localização das larvas propriamente ditas, que estariam concentradas ao redor dos ovos. Neyman propôs, então, um modelo alternativo, por ele batizado de "distribuição tipo A", para explicar o padrão de distribuição das larvas nas diversas porções do terreno. O novo modelo proposto foi confirmado pelos resultados experimentais e pelos testes de significância realizados.

A conclusão que Gillies extrai é que Neyman conseguiu explicar um resultado experimental inesperado a partir da metodologia popperiana de testar hipóteses e utilizar testes estatísticos para refutá-las. Com isso, as investigações de Neyman com relação à distribuição das larvas constituiria um exemplo bem sucedido dos méritos do falsificacionismo. Tendo exibido esses méritos, Gillies faz uso do mesmo exemplo para mostrar que o bayesianismo não conseguiria auxiliar Neyman a desvendar o significado dos indícios experimentais que refutaram a hipótese inicial.

Para investigar a distribuição das larvas, um cientista bayesiano formularia, inicialmente, um conjunto H_λ de hipóteses alternativas, em que $0 < \lambda < \infty$. Em seguida, esse cientista estimaria uma distribuição de probabilidades prévias $p(\lambda)$ para cada uma dessas hipóteses, que refletiria seu grau prévio de crença em cada uma delas.

As distribuições de probabilidades prévias seriam, em seguida, modificadas à luz dos indícios experimentais obtidos, de modo a obter a probabilidade posterior $p(\lambda/e)$.

Gillies considera difícil ver como o enfoque bayesiano poderia ter auxiliado na resolução do problema, isto é, o abandono da hipótese relativa ao modelo de Poisson pela hipótese da distribuição tipo A. O bayesianismo, prossegue Gillies, não faz muito além de permitir a condicionalização das probabilidades prévias das hipóteses alternativas à luz dos indícios recolhidos. Não auxilia, porém, a elaboração de uma nova hipótese. Isso seria uma demonstração clara de que o bayesianismo exige que o quadro teórico seja fixo, no caso, a hipótese de que as larvas se distribuíam conforme o modelo de Poisson.

Um bayesiano poderia alegar que seu enfoque funcionaria a contento se o conjunto inicial H_λ incluísse, além da hipótese do modelo de Poisson, também o modelo tipo A. A condicionalização mostraria, nessa situação, que a probabilidade posterior no primeiro caso decairia com relação à probabilidade prévia e, no segundo caso, aumentaria. Entretanto, a dificuldade com esse argumento é que, conforme narra o próprio Neyman, o modelo tipo A somente foi concebido após a refutação da hipótese original pelos indícios. Mesmo que H_λ incluísse todas as hipóteses para explicar o fenômeno a ser analisado, o que está longe de ser evidente, Gillies acredita que o emprego do método falsificacionista levaria à solução do problema muito mais rapidamente que o bayesianismo. O cientista bayesiano estaria ainda lutando para determinar a probabilidade posterior de suas hipóteses no momento em que o falsificacionista já teria formulado sua nova hipótese *h´* e resolvido a questão.

A conclusão de Gillies é que como o bayesianismo depende da fixidez do quadro teórico, os cientistas bayesianos são forçados a enfrentar uma escolha difícil. Eles devem, ao início de sua investigação, considerar o maior conjunto possível de hipóteses alternativas e torcer para que uma delas seja adequada; ou eles devem correr o risco de nunca chegar à hipótese correta para explicar os resultados experimentais obtidos. A dificuldade é uma consequência direta, insiste Gillies, da própria essência do bayesianismo, qual seja, a imposição de limitações inerentes às modificações nos graus de crença produzidas pela condicionalização bayesiana.

Antes de partir para o segundo exemplo de Gillies, é necessário tecer, a esta altura, um comentário em defesa do bayesianismo, mesmo se Howson e Urbach não tenham respondido diretamente ao autor. As conclusões de Gillies não parecem decorrer da linha de raciocínio por ele seguida. Em particular, as razões de o bayesianismo não propiciar o abandono de uma hipótese equivocada em favor de outra não parecem estar claras. No exemplo de Neyman, não haveria qualquer impedimento a que ele, ao atualizar a probabilidade prévia de sua hipótese original com base no Teorema de Bayes, chegasse à conclusão de que ela havia sido infirmada em tal grau que sua manutenção seria insustentável. Com isso, Neyman seria forçado a buscar hipóteses alternativas e, eventualmente, poderia chegar a formular a hipótese da distribuição tipo A. O fato de o bayesianismo trabalhar com as categorias de probabilidade posterior e prévia não impede que uma hipótese seja abandonada no caso de ser infirmada pelos indícios recolhidos. A questão é apenas saber quantas instâncias de infirmação teriam sido necessárias para que a hipótese original fosse abandonada, o que não significa, em absoluto, que o apego ao enfoque bayesiano impediria a formulação da nova hipótese.

O segundo exemplo empregado por Gillies para delinear os limites de aplicação do bayesianismo diz respeito às pesquisas de De Finetti acerca da questão da permutabiliade e sua relação com a chamada independência probabilística. Também aqui Gillies procura mostrar que o bayesianismo só é aplicado se o quadro teórico for fixo. Em um artigo de 1937 (1937, pp. 93-158), De Finetti se coloca a seguinte pergunta: “Por que estamos obrigados, na maioria dos problemas, a avaliar a probabilidade a partir da observação de uma frequência?”. De Finetti imagina o conhecido experimento de jogar uma moeda e registrar se o resultado obtido foi “cara” ou “coroa”. De Finnetti utiliza esse exemplo em função da generalidade das conclusões que acredita poder dele extrair com relação ao problema da indução. Em sua tentativa de mostrar os limites do bayesianismo, a estratégia de Gillies será a de circunscrever as conclusões que extrai De Finnetti, qualificando-as de não generalizáveis.

O primeiro passo de De Finetti é considerar uma sequência de resultados quando a moeda é lançada: $E_1, E_{2,} \dots, E_n, \dots$. Os resultados obtidos podem ser “cara” (C_i) ou “coroa” (Co_i). Em particular, $C_{n+1} =$ “cara” ocorre no lançamento $n + 1$. Seja agora e o registro de resultados após n lançamentos, em que o resultado “cara” ocorre r vezes. Ao aplicar o método de condicionalização bayesiana, é possível calcular $P(H_{n+1}/e)$, que corresponde à probabilidade posterior, dado e, de se obter “cara” no lançamento $n + 1$. Com base nesse cálculo, é possível mostrar que, satisfeitas determinadas condições, a probabilidade posterior $P(H_{n+1}/e)$ tende a $\frac{r}{n}$ quando n tende ao infinito. Se uma moeda for lançada um milhão de vezes e o resultado "cara" for observado aproximadamente a metade das vezes, nada é mais razoável do que considerar a frequência de $\frac{1}{2}$ como a probabilidade prévia do resultado "cara". Esse raciocínio tem a finalidade, para o bayesiano, de mostrar que, mesmo no caso de indivíduos com opiniões e atribuições de probabilidades prévias muito diferentes entre si, eles irão coincidir na atribuição de probabilidades posteriores se utilizarem a regra de condicionalização para atualizar suas respectivas probabilidades prévias. Dessa forma, De Finetti buscou deixar claro que, em situações similares ao lançamento de uma moeda, os bayesianos subjetivos irão sempre escolher probabilidades posteriores próximas à frequência do evento em exame, como é indicado intuitivamente pelo exemplo da moeda.

A partir dessas considerações iniciais, De Finetti dá um passo adicional e tenta relacionar dois conceitos: permutabilidade e independência. Ele faz isso por meio de um teorema específico. Ao relacionar esses dois conceitos, o objetivo de De Finetti é qualificar como sem sentido a ideia de probabilidade fixa e desconhecida e, com isso, desvelar um golpe no frequentismo, em favor do bayesianismo.

Um conjunto de eventos $X_{i1}, X_{i2}, \ldots, X_{in}, \ldots$ é permutável se sua distribuição conjunta for sempre a mesma, independentemente de como $i_1, i_2, \ldots, i_n$ sejam escolhidos. Em outras palavras, não importa a ordem em que os eventos sejam tomados, a distribuição final sempre será sempre a mesma. Imagine-se, por exemplo, o caso de dez lançamentos de uma moeda em que são registrados seis resultados “cara” e quatro resultados “coroa”. Não importa a ordem dos resultados, apenas que, ao final, foram obtidos seis “caras” e quatro “coroas”. Qualquer sequência é permutável (substituível) com as demais, desde que se tenham seis “caras” e quatro “coroas”.

No exemplo aludido, há duzentas e dez sequências distintas de obter o resultado seis “caras” e quatro “coroas”, resultado que é obtido a partir da fórmula matemática para o número de combinações possíveis para seis resultados “cara” em dez lançamentos (${}_{r}^{n}C = \frac{n!}{r!(r-n)!}$). O teorema de De Finetti prova, em poucas palavras, que a distribuição de probabilidades no caso de um conjunto de eventos permutáveis é uma combinação linear das distribuições de probabilidades de eventos independentes e igualmente prováveis. Isso quer dizer que a distribuição de probabilidades do resultado seis “caras” e quatro “coroas” constitui a combinação das distribuições de probabilidades de dez lançamentos independentes em que existem dois resultados (cara x coroa), os quais são igualmente prováveis (supõe-se que a moeda seja não viciada).

O teorema de De Finetti é importante não apenas em si mesmo, mas porque serviu de instrumento para a formulação de uma poderosa crítica aos frequentistas. Embora no exemplo dado acima tenha se feito uso de eventos cuja probabilidade é conhecida (lançamento de uma moeda não viciada), o teorema pode ser estendido a eventos cuja probabilidade é desconhecida de antemão. Imagine-se o caso de uma moeda viciada em que não se sabe a probabilidade de se obter o resultado "cara" ou o resultado "coroa". Um frequentista caracterizaria o conjunto de dez lançamentos dessa moeda como uma situação de eventos independentes em que a probabilidade do resultado "cara" ou do resultado "coroa" seria desconhecida. Ao generalizar a aplicação de seu teorema a situações como essa, De Finetti discorda dessa maneira de descrever tais casos. Para ele, a forma correta de descrever tais situações seria por meio da noção de eventos permutáveis. A ideia de probabilidades fixas desconhecidas seria desprovida de sentido. Nesse artigo de 1937, De Finetti afirma que, “no caso de uma moeda irregular e viciada, o analista não deve considerar que a irregularidade física da moeda exerce uma influência sobre a 'probabilidade desconhecida' de se obter 'cara' ou 'coroa', pois essa probabilidade desconhecida não pode ser definida”.

Como defensor do falsificacionisto objetivista, Gillies não pode concordar com essa conclusão. Ele tentará mostrar que a ideia de probabilidades fixas desconhecidas não apenas possui sentido, como também que constitui uma prática perfeitamente aceitável postular a existência de processos constituídos por eventos independentes constituídos por elas. Para isso, Gillies procurará mostrar que o teorema de De Finetti acerca da permutabilidade apenas funciona nos estreitos limites do quadro teórico de eventos que são independentes. Nessa situação específica (eventos independentes), é possível abrir mão da noção de probabilidades fixas independentes e avaliar o fenômeno por meio da noção de eventos permutáveis. Ao se lidar, porém, com eventos probabilisticamente dependentes, o teorema de De Finetti deixa de ser aplicável e a única forma de interpretar o fenômeno é por meio da noção de probabilidades fixas desconhecidas.

Uma vez mais, Gillies faz uso de um exemplo concreto para mostrar que o teorema de De Finetti tem aplicação restrita a uma situação em que o quadro teórico é fixo. Gillies escolhe um exemplo de sequência de eventos dependentes: o jogo “vermelho ou azul”, originalmente mencionado por Feller. A cada rodada do jogo há um número *s*, o qual é determinado pelos resultados anteriores. Imagine-se uma moeda não viciada que é lançada. Se o resultado é “cara”, o número *s* é alterado para $s' = s + 1$;

se o resultado é “coroa”, o número s é modificado para $s' = s - 1$. Se $s' \geq 0$, então o resultado da rodada é considerado “azul”. Se, por outro lado, $s' \leq 0$, o resultado é considerado “vermelho”. Apesar de o jogo estar baseado no lançamento de uma moeda não viciada e do fato de cada lançamento ser independente, o resultado final "vermelho" ou "azul" é altamente dependente. Espera-se que o jogo produza longas sequências de “azul”, embora não se espere que sejam produzidas longas sequências de “caras”. Se, ao iniciar o jogo, estabelece-se que $s = 0$, então verifica-se um desequilíbrio em favor do resultado “azul”. Isso é facilmente corrigível ao se estabelecer que $s = -1$, o que faz que os resultados “azul” e “vermelho” sejam simétricos. Mesmo assim, à medida que várias rodadas se sucedem, há uma grande probabilidade de que uma das duas cores venha a predominar sobre a outra. Um dos exemplos mencionados por Feller é de que se o jogo for jogado uma vez a cada segundo, durante um ano, há 70% de probabilidade de que uma das cores será dominante 73% das vezes, enquanto a outra somente aparecerá nos 27% restantes.

A questão que Gillies coloca após apresentar o jogo "vermelho ou azul" é se o enfoque bayesiano estaria em condições de interpretá-lo corretamente. Longe de se tratar de uma situação hipotética, o jogo "vermelho ou azul" pode ser utilizado para modelar fenômenos correntes, por exemplo, uma sequência de dias que são classificados como "chuvosos" ou "ensolarados". De fato, em um estudo mencionado por Gillies (2001), concluiu-se que a sequência dos dias poderia ser corretamente modelada dessa forma. Concluiu-se, ainda, que a probabilidade de um dia ser ensolarado, dado que o dia anterior havia sido ensolarado, era 0,75. De maneira similar, a probabilidade de um dia ser chuvoso, dado que tinha chovido na véspera, era 0,66.

Para interpretar uma sequência de resultados modelada com base no jogo "vermelho ou azul", um bayesiano iniciaria sua análise assumindo a permutabilidade dos resultados, conforme exige o teorema de De Finetti. Muito provavelmente, esse bayesiano iria assumir uma distribuição de probabilidades *a priori* uniforme, o que significa que a cada rodada seria atribuída a mesma probabilidade ao resultado "vermelho" ou "azul", isto é, $p(e_1) = p(e_2) = \cdots = p(n)$. Uma vez que, na prática do jogo, essas probabilidades não são idênticas, pois o resultado de cada rodada depende do resultado da rodada anterior, o cientista bayesiano terá grandes dificuldades em utilizar o Teorema de Bayes para atualizar suas probabilidades e interpretar corretamente o fenômeno.

Suponha-se uma sequência de setecentos resultados "azul", seguidos por dois resultados "vermelho". O bayesiano calculará a probabilidade de o próximo resultado ser "azul" com base no limite da frequência observada, o será equivalente a 0.996. Alguém familiarizado com a natureza do jogo, porém, estimaria a probabilidade de o próximo resultado ser "azul" de forma bem diferente. Essa pessoa saberia que, para que o resultado 701 tenha sido "vermelho", é necessário que, no lançamento n. 700, o valor de *s* fosse zero e o resultado do lançamento da moeda tenha sido "coroa". Com isso, *s* assumiu o valor de -1 no lançamento n. 701. Como o lançamento seguinte também foi "vermelho", sabemos que o resultado produzido também foi "coroa" e que, com isso, o valor de *s* passou a ser -2. Como *s* é igual a -2 no lançamento 703, sabemos que a probabilidade de se obter o resultado "azul" é zero. Por mais que fizesse uso do processo de condicionalização, um bayesiano dificilmente chegaria a esse resultado.

Os cálculos realizados com base na premissa da permutabilidade de eventos levam a resultados muito distantes da verdade, se os eventos forem dependentes. Isso ocorre, segundo Gillies, porque implicitamente é assumido que o processo em exame consiste de eventos independentes com probabilidades fixas e desconhecidas. Na medida em que esse quadro teórico não se verifica, a utilização do enfoque bayesiano leva a resultados inapropriados. Um falsificacionista, por outro lado, mesmo que considere inicialmente que a sequência de 702 resultados seja um conjunto de eventos independentes, será forçado a revisar essa hipótese à luz dos resultados experimentais obtidos. Ele abandonará a hipótese da independência e passará a trabalhar com a possibilidade de os eventos serem dependentes entre si e, com isso, chegará a uma solução satisfatória para a questão. Com o exemplo do jogo "azul ou vermelho", Gillies conclui sua crítica ao bayesianismo de que este apenas funciona a contento no âmbito de um quadro teórico fixo.

ELLIOT SOBER, PROBABILIDADES PRÉVIAS E O PROBLEMA DA SIMPLICIDADE

O terceiro ataque ao bayesianismo a ser apresentado neste capítulo foi elaborado por Eliott Sober em seu artigo *Bayesianism – Its Scope and Limits*, publicado no volume *Baye´s Theorem (Proceedings of the British Academy)* pela Oxford University Press, em 2002. Nesse artigo, Sober não toma partido no debate entre bayesianos e falsificacionistas. Sem deixar de reconhecer que o bayesianismo exerce um papel importante na análise do raciocínio científico - seja em sua vertente objetivista, seja em sua vertente subjetivista -, ele indica que há um conjunto de situações no âmbito da atividade dos cientistas que o enfoque bayesiano não consegue dar conta. Ao apresentar tais situações, Sober concentra-se, especificamente, em dois aspectos do bayesianismo: a questão das probabilidades prévias e a questão da simplicidade.

O tratamento dado às probabilidades prévias constitui, na avaliação dos críticos do bayesianismo, um de seus principais pontos fracos, uma vez que introduziria elementos de indesejado subjetivismo na análise raciocínio científico, por mais que Howson considere tal crítica desprovida de impacto, conforme se viu no capítulo anterior. O cerne da questão diz respeito a situações em que a atribuição de probabilidade prévia a uma hipótese é feita, única e exclusivamente, com base no grau de crença do pesquisador, independentemente de qualquer dado empírico.

Na análise da questão das probabilidades prévias, Sober apresenta, inicialmente, uma possível linha de defesa para o bayesianismo, baseada na ideia de verossimilhança. Essa linha de defesa é, a seguir, rompida por Sober, que mostra haver situações em que, apesar do recurso à noção de verossimilhança, a explicação bayesiana permanece insatisfatória, o que deixa claro haver limites intransponíveis para a utilidade desse enfoque como instrumento de análise.

Uma forma possível de fugir ao subjetivismo inerente à atribuição das probabilidades prévias é, de fato, o recurso à noção de verossimilhança para comparar hipóteses rivais. Nesse caso, as hipóteses rivais serão comparadas entre si, mas tal comparação será feita exclusivamente com base no impacto de determinado indício sobre cada uma delas, sem que seja necessário atribuir valores de probabilidades prévias para fins de comparação. Isso significa que o procedimento para comparar duas hipóteses rivais h_1 e h_2, à luz de determinado indício *e*, deve ser o de comparar o valor de seus respectivos graus de verossimilhança, isto é, $p(e/h_1)$ e $p(e/h_2)$. É possível, ainda, formular a conceito de "confirmação relativa" (*differential support*), que constitui a razão entre os respectivos graus de verossimilhança, isto é, $\frac{p(e/h_1)}{p(e/h_2)}$ (Forster e Sober, 2002). Sober ressalta, por fim, que as noções de verossimilhança e confirmação relativa não indicam ao cientista em qual hipótese ele deve acreditar, nem tampouco qual hipótese é provavelmente verdadeira; elas simplesmente indicam um procedimento para comparar os respectivos graus de confirmação das hipóteses em consideração.

É importante ter presente que os bayesianos nem sempre consideram adequada a utilização da noção de verossimilhança para defender o bayesianismo. De fato, os bayesianos apontam que o uso da verossimilhança pode levar a conclusões discutíveis, uma vez que hipóteses absurdas podem apresentar grau de verossimilhança máxima. É o caso, por exemplo, da hipótese segundo a qual um jogador retirou de um maço de cartas um seis de espadas precisamente porque um demônio maligno o fez retirar especificamente aquela carta. Se tomarmos, como hipótese rival, a hipótese de que o resultado se deve a um processo aleatório, a verossimilhança da primeira hipótese será maior do que a da segunda. Isso acontece, apontam os bayesianos, justamente porque os defensores do uso da noção de verossimilhança hesitam em atribuir probabilidades prévias às hipóteses que comparam. Se o fizessem, continuam, ficaria claro que não obstante tenha sido retirada a carta seis de espadas, a probabilidade posterior da primeira hipótese permanece baixa se comparada à probabilidade posterior da segunda hipótese.

De qualquer forma, o argumento de Sober se articula no sentido de mostrar que embora o uso da noção de verossimilhança seja imune à crítica de subjetivismo, uma vez que não são atribuídas probabilidades prévias às hipóteses em comparação, ele não está livre de problemas. O problema surge no caso das chamadas hipóteses compostas, isto é, hipóteses que são constituídas por disjunções mutuamente excludentes de hipóteses simples. Nesses casos, não é possível calcular a verossimilhança da hipótese composta. O exemplo utilizado por Sober é o seguinte: imagine-se que, diante do indício constituído por um indivíduo com os genes Aa, se deseja calcular a verossimilhança da hipótese de que sua mãe apresente os genes AA, o que em símbolos seria expresso por $P(F_{Aa}/M_{AA})$, em que F expressa a condição de “filho” e M a condição de “mãe”.

Segundo as leis de Mendel, para que o filho exiba os genes Aa, dado que sua mãe é AA, é necessário que seu pai seja aa ou Aa. Isso ocorre porque o filho herda, necessariamente, o gene A da mãe e que, portanto, somente poderá herdar o gene a do pai que, nessa situação, deve ser Aa ou aa. Nesse caso, a verossimilhança que queremos calcular será a soma de duas probabilidades: a primeira, a probabilidade de o filho ser Aa, dado que a mãe é AA e o pai é aa; a segunda, a probabilidade de o filho ser Aa, dado que a mãe é AA e o pai é Aa. Isso pode ser expresso da seguinte forma:

$$P(F_{Aa}/M_{AA}) \;=\; P(F_{Aa}/M_{AA}.P_{Aa}) \times P(P_{Aa}/M_{AA}) \;+\; P(F_{Aa}/M_{AA}.P_{aa}) \times P(P_{aa}/M_{AA})$$
$$= \frac{1}{2} \times P(P_{Aa}/M_{AA}) \;+\; 1 \;\times\; P(P_{aa}/M_{AA}) \;\;=\; \frac{1}{2} \times\; w_1 \;+ 1 \;\times\; w_2$$

Para calcular a verossimilhança da hipótese de a mãe ser AA, dado que seu filho é Aa, é necessário conhecer os valores dos fatores de ponderação w_1 e w_2, que constituem expressão do comportamento de acasalamento na espécie em questão (probabilidade de um macho *Aa* acasalar com uma fêmea, dado que essa fêmea é *AA;* probabilidade de um macho *aa* acasalar com uma fêmea, dado que essa fêmea é *AA*). O problema é que essas informações simplesmente não existem. Não há qualquer indício acerca de se o comportamento no acasalamento se dá ao acaso ou de forma associativa, nem em que proporção em um ou outro caso. Assim, na falta de qualquer indício, torna-se extremamente difícil – na verdade, impossível – especificar valores para os fatores de ponderação, a menos que isso seja feito exclusivamente com base em graus de crença do pesquisador.

Sober indica que alguns pesquisadores, conscientes da dificuldade de determinar valores para os fatores de ponderação, deliberadamente optam por escolher valores que maximizam a verossimilhança da hipótese composta. No caso específico, eles assumem que a fêmea *AA* sempre irá se acasalar com um macho *aa*, ignorando a possibilidade de o macho ser *Aa*. Trata-se de um procedimento artificial, que em lugar de permitir determinar a verossimilhança da hipótese original, apenas permite determinar a verossimilhança de uma de suas componentes como se fosse a hipótese original.

Em suma, o fato de que determinadas hipóteses compostas frequentemente não dispõem de graus de verossimilhança objetivos revela-se uma barreira intransponível, segundo Sober, para o bayesianismo. A tentativa de escapar do subjetivismo inerente à atribuição de probabilidades prévias levou Sober a examinar se o recurso à verossimilhança não seria uma alternativa em condições de salvar o projeto bayesiano dessa acusação. Tal tentativa revelou-se, porém, malograda, uma vez que, tanto como no caso das probabilidades prévias, também no caso da verossimilhança há situações – aquelas relativas a hipóteses compostas – que não permitem a atribuição de valores objetivos. Essa não é, porém, a única limitação apontada por Sober ao projeto bayesiano.

O segundo ponto levantado por Sober como limitação intransponível do bayesianismo diz respeito ao critério de simplicidade para comparar hipóteses rivais. No capítulo anterior, viu-se como Popper faz uso da noção de simplicidade para criticar o bayesianismo, assim como a resposta de Howson e Urbach. Ao tratar da questão da simplicidade, Sober quer mostrar que o bayesianismo é incapaz de justificar de forma adequada a atribuição de probabilidades prévias mais altas a hipóteses mais simples. Para isso, Sober desenvolve seu argumento com base em uma questão específica que não havia sido devidamente tocada no debate entre o falsificacionismo e bayesianismo, a saber, se e como os cientistas devem comparar os chamados modelos aninhados (*nested models*).

Dois modelos são considerados aninhados quando estão relacionados de tal forma que um deles é a extensão do outro. Em seu texto, Sober utiliza, como exemplo, dois modelos para explicar os traços genéticos de uma espécie. O primeiro modelo é classificado como N5, porque contém apenas cinco parâmetros ajusáveis; o segundo é classificado como N90, uma vez que apresenta noventa parâmetros ajustáveis. É uma consequência do cálculo de probabilidades que, se N90 decorre de N5, então a probabilidade de N5 é menor ou igual à de N90.

Quando dois modelos não são aninhados, não há qualquer dificuldade com a comparação entre eles a partir do critério de simplicidade. Não há, tampouco, qualquer impedimento a que ao modelo mais simples se atribua uma probabilidade prévia mais alta. O que fazer, porém, diante de dois modelos explicativos aninhados? Inicialmente, um bayesiano poderia indicar que não tem sentido considerar como rivais dois modelos aninhados, já que eles não são incompatíveis. Sober assevera que a insistência na comparação de modelos não aninhados não elide sua crítica ao bayesianismo, uma vez o bayesianismo é incapaz de justificar a atribuição de probabilidade prévia mais alta a um modelo mais simples, seja no caso de modelos aninhados, seja no caso de modelos não aninhados.

Imagine-se uma versão alternativa do modelo N90 que não mais seja dedutível de N5. Um tal modelo utilizaria diferentes parâmetros e poderia ser classificado como N90*. Que razão poderia ser apresentada para atribuir a N5 uma probabilidade prévia mais alta do que a N90*? O bayesianismo, assevera Sober, não tem qualquer resposta a essa pergunta. Seria possível, ele reconhece, tentar justificar a atribuição de probabilidades prévias a partir da noção de verossimilhança, mas nesse caso ressurgiriam as dificuldades já apontadas.

O bayesianismo possui apenas dois instrumentos para justificar a relevância epistêmica da noção de simplicidade: as probabilidades prévias e a noção de verossimilhança. Nenhum desses dois instrumentos consegue dar conta da tarefa, segundo Sober. Diante disso, a única alternativa de que dispõe o bayesiano, segundo Sober, é negar que a noção de simplicidade possua qualquer relevância. Ela seria apenas objeto da predileção dos cientistas por várias razões, mas não teria qualquer significado epistêmico.

O movimento bayesiano de tentar negar relevância à noção de simplicidade esbarra, porém, em duas dificuldades. A primeira é que, embora haja situações em que a noção de simplicidade é irrelevante do ponto de vista da indução, há outras em que ela se revela essencial. É o caso, por exemplo, da seleção de modelos alternativos na biologia e nas ciências sociais. A segunda dificuldade é que existe um modelo para explicar a indução, desenvolvido por H. Akaike, que mostra com clareza a importância da noção de simplicidade, entendida como número de parâmetros ajustáveis, para o grau de precisão de um modelo empregado para realizar predições. No modelo de Akaike, a simplicidade é um elemento crucial do ponto de vista da precisão do modelo.

Sober acredita ter exposto de forma terminal as limitações do bayesianismo a partir do exame que realiza das noções de verossimilhança e simplicidade. Para ele, os bayesianos não conseguem justificar a importância da noção de simplicidade, seja por meio da forma como atribuem probabilidades prévias, seja por meio da noção de verossimilhança. Quando dois modelos estão aninhados, é impossível que o modelo mais simples tenha probabilidade prévia mais alta que o mais complexo, pois isso violaria uma das consequências dos axiomas do cálculo de probabilidades. Para modelos não aninhados, não há qualquer impedimento a que se atribua maior probabilidade prévia ao modelo mais simples, mas nesse caso o bayesianismo não consegue aportar qualquer justificativa convincente. Sober nega que essa atribuição de probabilidades possa ser justificada com base no maior grau de crença do pesquisador no modelo mais simples. Para ele, isso é simplesmente fugir da questão, que é saber as razões para ser correto considerar mais provável um modelo mais simples. A tentativa de justificar essa atribuição com base na noção de verossimilhança também fracassa, dado a objeção estabelecida em situações de hipóteses compostas. Assim, vê-se que, independente das várias direções que seu argumento toma, a crítica de Sober ao bayesianismo está baseada, antes de tudo, na negação de que é possível justificar a indução a partir de probabilidades subjetivas.

DEBORAH MAYO, O PROBLEMA DE DUHEM E O BAYESIANISMO

As últimas observações críticas ao projeto bayesiano a serem analisadas neste capítulo dizem respeito a alegadas anomalias geradas quando ele é empregado para solucionar o problema de Duhem, isto é, a questão relativa a qual hipótese deve ser responsabilizada quando uma teoria é confrontada com uma instância de infirmação. Tais críticas foram analisadas por vários autores e expostas em sua essência por Deborah Mayo no texto *Response do Howson and Laudan* (Mayo, 1997).

Ao responder a Howson e Urbach, Mayo quer mostrar que o enfoque bayesiano não consegue explicar como os cientistas lidam com o problema de Duhem, isto é, como eles concluem qual hipótese devem culpar em caso de uma instância de infirmação. De forma subsidiária, Mayo também quer mostrar que para solucionar esse problema, os cientistas recorrem a métodos estatísticos para distinguir efeitos reais de artefatos, os mesmos métodos que o bayesianismo descarta ao analisar o raciocínio científico.

O cerne do ataque de Howson e Urbach ao falsificacionismo, com relação ao problema de Duhem, diz respeito aos testes de significância, especificamente os testes de Neyman-Parsons. Para Howson e Urbach, para solucionar o problema de Duhem e escolher qual hipótese deve descartar, o cientista deve dispor de algum instrumento que lhe permita saber se determinada hipótese é aceita, dado que é falsa, ou seja, $P(H_{aceito}/{\sim}H)$. Ocorre, porém, que o que o teste de Neyman-Parsons permite saber é a probabilidade de a probabilidade H ser falsa dado que é aceita, isto é, $P(H_{falsa}/H_{aceita})$. Essa última probabilidade condicional apenas informa que algumas hipóteses aceitas são, em realidade, falsas, mas não auxilia o investigador a saber de qual hipótese se trata. Essa última informação é dada, porém, pela primeira probabilidade condicional.

O primeiro movimento de Mayo em defesa dos testes de Neyman-Parsons é considerar injusta a crítica de Howson e Urbach. Mayo nega que os cientistas confundam as duas probabilidades condicionais. Nesse primeiro momento, ela afirma (p. 324) que a crítica de Howson provavelmente se deve à falta de conhecimento daquele autor sobre a aplicação de métodos estatísticos no dia a dia dos laboratórios.

Em um segundo momento, Mayo também rebate a crítica adicional de Howson de que os testes de Neyman-Parsons levariam a resultados equivocados. Na verdade, a afirmativa de Howson está baseada na constatação de que tais testes não se prestam a fornecer a probabilidade posterior de determinada hipótese. Em sua resposta, Mayo procurará mostrar que os testes de Neyman-Parsons e o enfoque bayesiano têm propósitos fundamentalmente opostos e, por essa razão, podem chegar a resultados dissimilares. Assim, o enfoque falsificacionista não levaria a resultados equivocados, mas simplesmente a resultados distintos, dada a diferença de propósitos envolvida, com relação ao bayesianismo.

Os testes de Neyman-Parsons, segundo Deborah Mayo, não guardam qualquer relação com as probabilidades prévias das hipóteses. Na verdade, eles existem para satisfazer determinados métodos estatísticos com vistas a permitir a realização de inferências quanto à verdade ou falsidade de determinada hipótese. Assim, tanto a utilização de probabilidades prévias quanto a determinação de probabilidades posteriores, são deliberadamente evitadas em vista dos objetivos que os testes procuram atingir. Dessa forma, prossegue Mayo, não faz sentido que Howson recrimine os testes de Neyman-Parsons por não realizar algo que nunca foi o seu propósito.

Mas Mayo não se contenta em defender o enfoque falsificacionista. Ela vai além e procura mostrar que o bayesianismo está longe de oferecer aquilo que promete. Howon afirma, segundo Mayo, que para solucionar o problema de Duhem, o cientista necessita saber a probabilidade condicional de a hipótese *h* ser falsa, dado que é aceita. A pergunta que se coloca, então, é se o bayesianismo é capaz de fornecer tal probabilidade condicional.

Suponha-se um newtoniano incorrigível que aceita a hipótese *A´*: o efeito de deflexão da luz é devido a um fator *N* qualquer (por exemplo, um efeito próprio das lentes de observação) que permite salvar as leis de Newton da refutação. Sabe-se, hoje, em função da Teoria da Relatividade, que *A´* é falsa. Nessas condições, o que interessa a um cientista confrontado com determinada instância de refutação é probabilidade condicional de *A´* ser falsa, dado que é aceita. Ocorre que tal probabilidade condicional tampouco é fornecida pelo bayesianismo, que apenas pode fornecer a probabilidade subjetiva para o agente de *A´* ser falsa, dado que é aceita. Em outras palavras, o bayesianismo não permite avaliar a veracidade ou falsidade da hipótese, dado que ela é aceita. Ele apenas fornece o grau subjetivo posterior de crença de que a hipótese é verdadeira ou falsa, dado que é aceita. Para um cientista adepto da estatística do erro, uma hipótese será conservada e considerada verdadeira se as margens de erro relacionadas aos testes a que foi submetida tiverem sido excluídos. Isso não é necessário, porém, para que a hipótese seja objeto de crença em sua veracidade.

O cerne da questão enfatizado por Mayo, e que será aprofundado no próximo capítulo, é que a estatística do erro não se importa diretamente com o grau de crença do cientista em determinada hipótese. O que importa para os filósofos que compartilham do enfoque de Mayo é saber as razões que permitem justificar a crença na veracidade ou falsidade de determinada hipótese de maneira objetiva. O bayesianismo, por sua vez, não está preocupado com a veracidade da crença, mas sim se tal crença é alcançada segundo um processo que seja consistente e coerente. Mayo acredita, portanto, que o bayesianismo não dá ao cientista aquilo que ele mais deseja, isto é, condições de posicionar-se quanto à veracidade ou falsidade de uma hipótese com base em critérios objetivos. Essa é, em poucas palavras, a grande diferença da estatística do erro para o bayesianismo, conforme será visto no próximo capítulo.

Com vistas a tornar mais clara essa diferença de enfoque entre o bayesianismo e a estatística do erro, Mayo fornece um exemplo em que contrasta os dois enfoques e aponta para as deficiências, segundo ela insuperáveis, do bayesianismo. O exemplo pode ser apresentado de forma resumida da seguinte forma: imagine-se determinada população apresenta determinada deficiência com a probabilidade de 0.999. Qualquer indivíduo escolhido aleatoriamente dessa população apresenta, portanto, uma probabilidade prévia igual a 0.999 de exibir a referida deficiência (por exemplo, a incapacidade para ser aprovado(a) em um exame competitivo as melhores universidades do país). Considere-se, diante dessa situação hipotética, duas hipóteses alternativas:

H: o (a) estudante não apresenta a deficiência;

$H´$: o (a) estudante apresenta a deficiência.

Há dois resultados passíveis de observação nesse caso: e (aprovação no exame) e $e´$ (não aprovação no exame). Suponha-se, ainda, que os estudantes que não apresentam a deficiência indicada são aprovados no exame com probabilidade prévia igual a 1; os estudantes deficientes são aprovados no exame com uma probabilidade prévia igual a 5%. Com isso, é possível calcular a probabilidade posterior temos que $P(H/e) \ = 1$

Com base nas probabilidades prévias e probabilidades posteriores, Mayo agora quer mostrar que o bayesianismo leva a um resultado que tem grandes chances de ser falso. A candidata Mary foi aprovada no exame competitivo para a universidade. Diante desse fato, qual das hipóteses deveria um cientista racionalmente escolher, H ou $H´$?

Segundo a leitura que Mayo faz de Howson, um cientista bayesiano estaria obrigado a escolher a segunda hipótese, pois a probabilidade posterior de H' será necessariamente maior do que a probabilidade posterior de H. O fato de a estudante ter sido aprovada terá um impacto muito maior em H' do que em H. De fato, com os dados estipulados, tem-se que $P(H'/e) = 0.95$. O resultado observado confirma muito mais a hipótese de que a estudante apresenta a deficiência do que a hipótese contrária. Isso acontece porque a estudante foi escolhida em uma área em que um indivíduo escolhido aleatoriamente tem probabilidade prévia igual a 0.999 de não estar em condições de ser aprovado no exame.

Mayo esclarece que o problema com o enfoque bayesiano no exemplo acima é que em momento algum foi considerada a possibilidade de H ser rejeitada erroneamente. E isso é o que interessaria aos cientistas, pois o fato de Mary ter sido aprovada no exame não garante que ela pertencia ao grupo dos despreparados, por mais que essa hipótese tenha sido confirmada em grau muito maior do que a sua rival. Muito provavelmente ela pertence ao grupo dos preparados e a hipótese H foi rejeitada erroneamente pela aplicação do enfoque bayesiano.

O cientista devoto do emprego dos métodos estatísticos não rejeitará H ao observar *e*, nem tampouco acreditará que esse indício é razão para aceitar H'. Caso a estudante Mary tivesse sido selecionada ao acaso de outra amostra populacional, em que as probabilidades prévias relativas à deficiência fossem distintas – por exemplo, em que a probabilidade prévia de ter a deficiência fosse 0.5 e, não, 0.9 –, então o impacto da aprovação sobre a hipótese H' seria muito menor. Com isso, o mesmo resultado, que em determinada situação levaria o analista a concluir, segundo o enfoque bayesiano, pelo despreparo de Mary, no segundo exemplo levaria à conclusão oposta. Para Howson e Urbach, isso seria uma consequência natural da modificação das probabilidades prévias e um resultado coerente do ponto de vista bayesiano.

Para Mayo, o problema com o bayesianismo é que embora ele seja internamente consistente, ele não dá ao cientista aquilo que ele de fato necessita. No caso de Mary, o que importa é garantir que o indício seja interpretado de forma correta, o que significa excluir qualquer erro de interpretação relativo ao desempenho da candidata. O objetivo a ser buscado, assim, é avaliar a possibilidade hipótese de Mary ser despreparada ser erroneamente tomada. E isso é, precisamente, o que a estatística do erro se propõe a fazer, conforme se verá no próximo capítulo.

CONCLUSÃO

Neste capítulo, foram apresentadas as alegadas deficiências que alguns dos críticos do bayesianismo identificam nesse enfoque. Os pontos levantados por Gillies acerca da rigidez teórica e por Sober apontaram para as dificuldades de aplicação do enfoque em situações específicas. Para além dessas críticas, é possível ver, a esta altura, que as consequências de certo grau de subjetividade inerente à utilização do Teorema de Bayes constituem, possivelmente, o grande pomo da discórdia entre o bayesianismo e o falsificacionismo. A utilização de probabilidades subjetivas e a aplicação do Teorema de Bayes podem levar a resultados que não são necessariamente verdadeiros, nem tampouco coincidentes com a aplicação do enfoque falsificacionista. Mas a pretensão de realizar inferências que levem a conclusões verdadeiras não é, possivelemente, o objetivo maior do bayesianismo e, sim, a realização de inferências válidas, baseadas na atualização das probabilidades por meio do mecanismo de condicionalização bayesiana. Para Howson e Urbach, o bayesianismo fornece um mecanismo confiável para a realização de induções. Tal mecanismo se funda na atualização dos graus de probabilidade prévia por meio da aplicação do Teorema de Bayes. Nem consistência das hipóteses e dos indícios, nem forma como estes foram recolhidos interessam, em princípio, ao bayesiano. Este crê, por outro lado, que se a atividade científica for realizada de forma conscienciosa, o acúmulo de indícios levará a resultados que serão verdadeiros.

No próximo capítulo, será brevemente apresentado o enfoque da estatística do erro, o qual tem em Deborah Mayo sua principal defensora. Tal enfoque, conforme se verá, constitui uma tentativa de revigoramento do racionalismo popperiano e uma resposta poderosa ao domínio do enfoque bayesiano durante um longo período nos departamentos de filosofia anglo-saxões, servindo de contraponto crítico a este e permitindo uma melhor compreensão da abordagem de Howson e Urbach.

4. A ESTATÍSTICA DO ERRO E A VOLTA DO FALSIFICACIONISMO

Em um artigo publicado em 1997, Deborah Mayo afirma que o bayesianismo tem sido o enfoque dominante na Filosofia da Ciência há algumas décadas. Como resultado dessa situação, aspectos essenciais da atividade científica foram negligenciados ou sua interpretação se mostrou profundamente equivocada. Para ela, o enfoque bayesiano estaria longe de interpretar corretamente indícios experimentais, o que lançaria muitas dúvidas acerca das inferências efetuadas a partir do Teorema de Bayes como uma correta interpretação do raciocínio científico.

É em resposta a essa situação que Mayo apresenta a chamada Estatística do Erro, programa de pesquisa que combina princípios do falsificacionismo popperiano com a aplicação rigorosa de métodos estatísticos para avaliar resultados experimentais. De fato, o objetivo principal da Estatística do Erro, conforme Deborah Mayo deixa claro já no primeiro capítulo de *Error and the Growth of Experimental Knowledge*, é submeter hipóteses a testes tão rigorosos quanto possível, com vistas a expurgar qualquer erro experimental que possa levar a inferências equivocadas.

O debate entre a Estatística do Erro e o bayesianismo constitui a mais recente edição de um embate que há tempos divide a Filosofia da Ciência entre interpretações do raciocínio científico que levam em conta elementos subjetivos ou psicossociais da atividade científica e interpretações que se pretendem mais objetivas, fundadas na noção de erro, nos testes de significância e outros métodos estatísticos. Em que consiste, exatamente, a Estatística do Erro e em quais pontos ela se distancia do bayesianismo? Este é o tema do último capítulo desta dissertação que, embora não tenha por escopo analisar todos os detalhes desse enfoque que se contrapõe ao bayesianismo, busca identificar seus traços principais segundo alguns escritos de Deborah Mayo, especialmente, mas não exclusivamente, *Error and the Growth of Experimental Knowledge*, de 1996.

A ESTATÍSTICA DO ERRO

Na análise do raciocínio científico por meio da aplicação de métodos probabilísticos, duas correntes principais podem ser identificadas (Pearson, 1950, p. 228): a primeira, chamada de abordagem indicial (*evidential-relation view*, E-R); a segunda, chamada de Estatística do Erro (*error statistical view*). A diferença entre as duas abordagens reflete diferenças fundamentais acerca de como considerações probabilísticas devem ser empregadas na realização de inferências.

A abordagem indicial surgiu como resposta à busca por uma lógica específica para a indução. Seu foco principal é mensurar o impacto dos indícios nas hipóteses ou teorias científicas. É nesse contexto que se encaixa o projeto bayesiano. Conforme se viu nos capítulos precedentes, o bayesianismo mede o impacto do indício e sobre a hipótese H por meio da aplicação do Teorema de Bayes, o que permite chegar à probabilidade condicional $P(H/e)$, a qual expressa a probabilidade posterior da hipótese em análise. Além disso, a aplicação do teorema exige a atribuição de probabilidade prévia à hipótese H. A probabilidade prévia deve refletir o grau de crença do investigador na hipótese H, segundo a vertente subjetiva do bayesianismo.

Em contraste com a abordagem indicial e o bayesianismo, os métodos e modelos clássicos da estatística, como os testes de significância e os intervalos de confiança, constituem exemplos da chamada Estatística do Erro. Tais métodos e modelos negam a possibilidade de atribuição de probabilidades prévias, a menos que estejam baseadas em frequências objetivas. A atribuição de probabilidades, por sua vez, é um instrumento para escrutinar e investigar o processo experimental. Todo processo experimental é caracterizado por margens de erro que podem ser expressas por probabilidades, as chamadas probabilidades de erro.

A Estatística do Erro apresenta duas dimensões complementares: uma filosófica; outra metodológica (Mayo e Spanos, 2010). A dimensão filosófica diz respeito a um enfoque próprio para a Filosofia da Ciência, que se baseia no papel do cálculo das probabilidades para realizar inferências. A dimensão metodológica, por sua vez, diz respeito ao emprego de um conjunto de métodos e procedimentos estatísticos, sua interpretação e justificação. Em última instância, a Estatística do Erro busca oferecer respostas a duas questões, que constituem seu foco principal: a) Como obter conhecimento confiável sobre o mundo, apesar da existência de incertezas e erros?; e b) Qual o papel das probabilidades na realização de inferências?

Na prática do dia a dia, há várias situações em que os métodos contemplados pela Estatística do Erro são empregados. É o caso, por exemplo, de pesquisas de opinião em que se procura inferir, com margem de p% a mais ou a menos, a proporção de um grupo que, provavelmente, irá votar no candidato X. Outros exemplos são relatórios sobre valores estatisticamente relevantes das diferenças entre um grupo tratado com um medicamento e seu respectivo grupo de controle; e análises de informação em experimentos em Física com o propósito de distinguir efeitos reais de artefatos. Em todos esses exemplos, a noção de erro experimental é essencial. Apesar desse uso difundido, a influência bayesiana fez, segundo Mayo, que por muito tempo os métodos aludidos tenham sido considerados inválidos pelos filósofos da ciência.

É preciso fazer um esclarecimento acerca do exato âmbito da Estatística do Erro. Embora esse enfoque faça uso de métodos típicos da estatística clássica, inclusive testes de significância e de Neyman-Pearson, ela o faz com vistas a constituir um enfoque novo. Como se verá mais adiante, esses métodos não constituem um fim em si mesmo. A cada experimento, os testes são empregados para permitir que uma hipótese seja testada da forma mais severa possível. A aplicação de testes de significância, segundo a Estatística do Erro, não se presta à aceitação ou descarte automático de hipóteses. Cada aplicação do teste deve ser revista à luz das especificidades do experimento em questão, tendo em vista o objetivo maior de realizar inferências confiáveis. O cientista que emprega a Estatística do Erro em sua prática laboratorial será levado a realizar questionamentos que vão muito além da mera aplicação do receituário recomendado nos manuais de estatística.

A Estatística do Erro procura desenvolver instrumentos para lidar com as inevitáveis limitações e erros próprios da atividade experimental. Diferentemente do bayesianismo, ela não propõe, nem dispõe de um mecanismo automático para avaliar o impacto de indícios em hipóteses. Os métodos estatísticos são empregados pelos investigadores para evitar que suas conclusões sejam equivocadas. Eles permitem aos cientistas desenvolver estratégias para coletar e modelar indícios, verificar a ocorrência de uma variedade de erros e, enfim, realizar inferências com o máximo grau de certeza possível.

Há dois aspectos da Estatística do Erro que são especificamente ressaltados por Mayo como virtudes em comparação com o bayesianismo (Mayo, 1997, p. 199). Em primeiro lugar, em lugar de fazer uso direto de indícios, a Estatística do Erro chama a atenção para a maneira segundo a qual estes foram recolhidos e modelados. Com frequência, será necessário realizar inferências específicas para que se possa chegar a indícios corretos do ponto de vista experimental (por exemplo, a necessidade de padronizar uma série de medidas para poder interpretá-las corretamente). Em segundo lugar, a Estatística do Erro não considera que a realização de inferências científicas possa ser levada a cabo através da aplicação de uma fórmula, como o Teorema de Bayes.

A Estatística do Erro constitui uma moldura investigativa (*a framework of inquiry*) para avalizar a realização de inferências. Não é possível simplesmente confrontar o cientista adepto da Estatística do Erro com determinado conjunto de indícios e esperar que ele esteja em condições de indicar, de imediato, se e qual hipótese é confirmada por esse conjunto. A moldura investigativa que propõe a Estatística do Erro exige que o procedimento para responder a tais perguntas seja fragmentado em uma série de etapas (*piecemeal approach*). Para se chegar à conclusão desejada, por exemplo, se os indícios confirmam ou não a hipótese, há uma sequência de passos a ser observada. Esses passos dizem respeito à realização de testes e identificação de possíveis erros desde a concepção de experimentos até sua efetiva realização, passando pela geração e modelagem de indícios.

Assim, de acordo com esse enfoque, uma investigação científica deve ser dividida em hipóteses separadas, as quais devem ser objeto de investigações independentes. Tais investigações independentes podem dizer respeito, por exemplo, a estimativas acerca de valores previstos pela teoria ou a testes hipotéticos dos mesmos. Elas correspondem a questões expressas em termos de erros-padrão de parâmetros, causas, efeitos acidentais e premissas envolvidas na investigação de outros erros. Deve-se identificar determinado erro e, a partir dele, inferir qual hipótese ou qual aspecto de uma hipótese é experimentalmente avalizado. Os modelos experimentais constituem o elemento de ligação entre modelos primários (a teoria) e os indícios modelados conforme o experimento.

Há, portanto, diferentes situações em que a utilização dos métodos estatísticos pode se dar. Uma primeira situação é quando esses métodos são empregados para analisar inferências a partir de hipóteses experimentais. Uma segunda situação é empregá-los para relacionar as inferências obtidas das hipóteses experimentais às previsões da teoria. Uma terceira situação diz respeito, por fim, à análise da transformação dos indícios de seu estado original em indícios modelados, isto é, que estão em condições de serem utilizados nas inferências a partir das hipóteses experimentais. A inferência de uma hipótese a partir de indícios pode ser descartada em função de um erro identificado na hipótese experimental ou porque os indícios obtidos não satisfazem ao que era previsto por essa mesma hipótese experimental.

TESTAR HIPÓTESES DE FORMA SEVERA, ARGUMENTAR A PARTIR DO ERRO

As considerações acima permitem entender porque, para Mayo (1997, p. 200), não é possível aspirar realisticamente às virtudes que o bayesianismo prega. No mundo real, os cientistas são confrontados com situações em que os indícios são inexatos, incompletos ou em que fatores externos não foram controlados ou não são controláveis. Inferências realizadas sem qualquer tipo de controle levarão inevitavelmente a resultados equivocados. Além disso, os cientistas não têm à sua disposição uma lista exaustiva de hipóteses e suas respectivas probabilidades prévias. Diferentemente do que propõe o bayesianismo, uma hipótese não deve ser aceita porque os indivíduos creem nela firmemente, nem é possível esperar até que eventuais desavenças acerca do poder de confirmação ou infirmação de determinado indício sejam decididas.

Ao utilizar a Estatística do Erro, o cientista pode descobrir muito acerca do mundo sem precisar da enorme quantidade de informação que, alegadamente, o bayesianismo exige, como a atribuição exaustiva de probabilidades prévias a hipóteses rivais. Mesmo em situações em que a condução de experimentos exaustivos não seja possível, Mayo indica que sempre será possível recorrer a experimentos contrafactuais ou realizar experimentos parciais de forma sequencial.

Uma questão que se coloca, a esta altura, é saber como é possível aprender com os erros. De forma simplificada, a busca pela identificação de erros nas várias etapas de uma investigação científica leva à elaboração de um conjunto de critérios eficientes para detectá-los. Se a despeito de todos os esforços empreendidos nenhum erro for encontrado, então o cientista tem excelentes motivos para acreditar que, de fato, nenhum erro está presente. Tal conclusão é resultado do emprego de procedimentos para testar de forma estrita a existência de determinado erro e mensurar seu valor (*severely probing error*).

Uma estratégia complementar identificada por Mayo (1997, p. 203) para se aprender com os erros é argumentar a partir deles (*arguing from error*). Após identificar e conhecer o suficiente sobre determinados erros próprios da prática experimental, o cientista está em condições de montar um experimento capaz de demonstrar a presença de um erro específico, caso este exista.

Um erro é considerado ausente na medida em que determinado procedimento investigativo, que pode ser composto por um ou vários experimentos, tem uma probabilidade alta de identificar esse erro se ele de fato estiver presente e, nessas condições, não o detecta. Do ponto de vista de uma hipótese H que se quer testar, o raciocínio proposto pela Estatística do Erro funciona da seguinte forma: indícios conformes à hipótese H indicam que ela é correta apenas e na exata medida em que tais indícios são o resultado de um procedimento experimental com alta probabilidade de produzir resultados não tão de acordo com H, caso esta seja falsa. Assim, os indícios indicam que H é correta se essa hipótese for submetida a um teste severo, no qual ela com alta probabilidade não passaria, se fosse falsa.

A fim de ilustrar o enfoque preconizado pela Estatística do Erro, tome-se uma hipótese H que estabelece que determinado efeito é real. A hipótese H estabelece, por exemplo, que existe uma correlação entre um gene específico e determinado tipo de câncer em uma população. Suponha, agora, que testes experimentais indicam a veracidade da hipótese H. Uma das conclusões passíveis de serem extraídas da veracidade de H é que a incidência do referido gene entre os pacientes que tem câncer não deverá variar da mesma maneira que se esperaria se essa correlação fosse aleatória. Nesse contexto, pode-se considerar a postulação de H como bem sucedida se ela for significativa do ponto de vista estatístico de forma distinta do que se poderia esperar se a relação entre o gene e a doença fosse aleatória. No caso de os testes indicarem que H é falsa, então resultados significativos estatisticamente acima de um nível determinado α , correspondente à correlação aleatória, não ocorrerão ou serão raros.

A noção de severidade se refere sempre a uma inferência em particular ou a uma hipótese que foi testada. Um experimento pode ser severo para testar uma hipótese, mas não outra. Nos parágrafos precedentes, a noção de severidade foi identificada com alta probabilidade de detectar o erro, se esse existir. Não há, porém, necessidade de que a noção de severidade seja construída com relação a qualquer cálculo estatístico. Com frequência, argumentos a partir do erro são construídos em avaliações qualitativas da noção de severidade.

Para testar de forma estrita a existência de determinado erro, o cientista deve subdividir a investigação que pretende levar a cabo em uma série de questões menores e relacionadas entre si. A cada questão corresponde uma hipótese parcial sobre a existência de um erro particular. A afirmação de que a hipótese H é verdadeira ou falsa dependerá da identificação de erros específicos. Por exemplo, se H afirma que determinado erro está ausente, H será falsa se esse erro estiver presente. Se H é verdadeira se um parâmetro qualquer for maior que c, então H será falsa se esse parâmetro for menor que *c*. Se *H* afirma que um fator *x* é responsável por ao menos *p%* de um efeito, sua negativa afirmará que esse fator é responsável por menos do que *p%*. Com isso permite-se que cada uma das questões particulares em que determinada investigação foi subdividida seja examinada de forma isolada.

Ao avaliar se uma pessoa tem dor de garganta causada pelo vírus *x*, o cientista adepto da Estatística do Erro deverá realizar um ou vários testes experimentais para discriminar o vírus *x* de outras possíveis causas. Ele trabalha com a dicotomia vírus *x versus* outras causas. Ele não está preocupado em atribuir probabilidades prévias às várias hipóteses que poderiam explicar a dor de garganta. No bayesianismo, por outro lado, avaliar uma hipótese singular é, necessariamente, compará-la com todas as demais que podem explicar o fenômeno investigado. Isso ocorre porque o bayesianismo trabalha com noção de um valor fixo de probabilidade total, igual a um, que deve ser distribuído entre as várias hipóteses rivais.

INFERÊNCIAS ILEGÍTIMAS?

Conforme já mencionado, o debate entre bayesianismo e Estatística do Erro constitui um dos eixos centrais ao redor do qual se desenvolveu a Filosofia da Ciência nas últimas décadas. Não é de estranhar, assim, que após a formulação da Estatística do Erro, nos anos noventa, Colin Howson tenha se apressado em defender o bayesianismo e, ao mesmo tempo, denunciar os "erros" da Estatística do Erro (Howson, 1997). Nesta seção, pretende-se comparar como uma mesma situação é interpretada alternativamente pelo bayesianismo e pela Estatística do Erro. Com isso, as diferenças entre um e outro enfoque ficarão mais claras e será possível ter uma ideia mais exata das virtudes e limitações tanto em um caso quanto em outro.

Ao criticar a Estatística do Erro, Howson ataca especificamente a estratégia de se argumentar a partir do erro (*arguing from error*) descrita acima. Vale a pena descrever todo o raciocínio de Howson, bem com a defesa de Mayo. Em poucas palavras, a estratégia de argumentar a partir do erro pode ser descrita, para Howson, da seguinte forma resumida (Howson, 1997): *e* constitui uma boa instância de confirmação de *H* na medida em que *H* passou por um teste severo relacionado a *e*. Para mostrar quão equívoco pode ser argumentar a partir do erro, Howson utilizará o exemplo de uma hipótese *H* que é confirmada por um teste severo que produziu o indício *e*. Não obstante, intuitivamente reconhece-se que *e* não confirma *H*, o que mostra que a conclusão que se extrai da argumentação a partir do erro é inválida.

O exemplo de que faz uso Howson é o mesmo do capítulo anterior, que será objeto de uma resposta mais desenvolvida de Mayo para mostrar aquilo em que consiste a Estatística do Erro em sua essência. O exemplo envolve uma população que está sujeita a uma doença qualquer com baixíssima incidência de *p%*. Uma pessoa escolhida aleatoriamente dessa população terá probabilidade prévia p de ter a referida doença. A hipótese *H* é de que a doença está presente.

Se uma pessoa escolhida exibe sintomas típicos da doença, temos, em princípio, um caso de um indício *e* que confirma a hipótese *H*. Para o filósofo adepto da Estatística do Erro, se o indício *e* foi obtido a partir de um teste suficientemente severo, então é possível inferir que a hipótese é verdadeira. Isso seria, todavia, um equívoco para Howson. O problema é que o indício *e*, apesar de ter sido obtido por um teste rigoroso, não seria capaz de tornar alta a probabilidade posterior da hipótese *H*. A hipótese *H* apenas poderia ser tomada como verdadeira após o acúmulo de uma quantidade suficiente de indícios e, com isso, sua probabilidade posterior viesse a se tornar alta. Howson acredita, assim, que há boas razões para acreditar que *H* é verdadeira apenas se sua probabilidade posterior for alta, à luz dos indícios acumulados. A Estatística do Erro, por outro lado, tornaria possível considerar *H* como verdadeira apenas com um indício, desde que, na obtenção desse indício, *H* tenha passado por um teste rigoroso.

É possível colocar o argumento de Howson de forma mais esquemática, da seguinte maneira (Mayo, 1997, p.207):

(1) Um resultado e (exibição de determinado sintoma da doença) não rejeita H (doença presente), ou seja, e confirma H. Segue-se que e rejeita a hipótese alternativa *J* (doença ausente);

(2) A hipótese H passou por um teste severo e pode, portanto, ser considerada verdadeira conforme a estratégia de argumentar a partir do erro;

(3) A doença em questão é tão rara na população que compõe o espaço amostral do qual o paciente foi selecionado que probabilidade posterior de *H*, isto é, $P(H/e)$ continua baixa. A probabilidade posterior de *J*, isto é, $P(J/e)$, por sua vez, permanece alta.

(4) Intuitivamente, *H* não deve ser considerada verdadeira, mas, sim, *J*.

(5) Consequentemente, ao argumentar a partir do erro, como feito em (2), deve ser considerada inválido.

Diante desse argumento, compete a Mayo apresentar a sua defesa da argumentação a partir do erro (Mayo, 1997, p. 207). Isso será feito em duas etapas: inicialmente, ela discordará da premissa (2) de Howson, que constitui uma interpretação equivocada daquilo que Mayo entende ser a argumentação a partir do erro. Em seguida, Mayo também mostrará que a premissa (4) de Howson não é válida no contexto da Estatística do Erro, pois está formulada no contexto do arcabouço bayesiano.

Analise-se a premissa (2). De fato, alguns testes de significância, se tomados sem qualquer consideração adicional, poderão levar à conclusão de que *H* é verdadeira. A questão para Mayo é saber se esses testes seriam considerados severos do ponto de vista da Estatística do Erro. Mayo nega que esse seja caso: o cientista adepto da Estatística do Erro não aplica testes de significância para aceitar ou rejeitar hipóteses automaticamente. Em lugar disso, o cientista deve inferir as hipóteses que passaram por testes severos. Ao mesmo tempo, saber se um teste é ou não severo exige que o cientista tenha clareza quanto ao tipo e extensão do erro que está sendo investigado. Embora não haja uma "receita" para levar a cabo esses procedimentos, Mayo acredita ser possível articular de forma sistemática um conjunto de regras "metaestasísticas", as quais permitiriam evitar rejeições ou aceitações equivocadas de hipóteses em situações de teste. Assim, Mayo acredita que a premissa (2) de Howson é equivocada, pois ele simplesmente ignora todas as considerações relativas à aplicação de testes pela Estatística do Erro. Isso porque um teste que não rejeite a hipótese *H* não deve levar automaticamente à conclusão de que essa hipótese é verdadeira, como ele indica.

No exemplo específico de Howson, o resultado *e* constituiria um indício pobre de que a doença em exame está presente acima de um nível *d* se esse indício também tiver grande probabilidade de ser verificado em níveis inferiores a *d*. Nesse caso, o indício *e* não constituirá um bom indicativo de que *H* é verdadeira (doença presente), pois o limite a partir do qual *e* é detectável é muito baixo para garantir, com alguma certeza, a presença da doença. De forma similar, o resultado *e* constitui um bom indício da presença da doença acima do nível *d* caso seja improvável que seja verificado em níveis inferiores a *d*. Nessa segunda situação, poder-se-ia afirmar que *e* daria boas razões para o cientista tomar *H* como verdadeira, pois apenas seria identificável no limite mínimo para que a doença estivesse presente.

Mayo nega que a premissa (2) do argumento de Howson possa ser sustentada. Se um teste não leva à rejeição de *H*, isso não significa que a hipótese possa ser considerada verdadeira, pois deve se verificar se existem alternativas a *H* que levem ao mesmo resultado. Assim, se o resultado *e* indicar que um indivíduo selecionado ao acaso exibe sintomas doença, é necessário investigar, por exemplo, se o teste experimental faz distinção entre tipos malignos e benignos da doença. Se não fizer, então *e* não poderá ser utilizado para afirmar a veracidade de *H* (doença presente).

Passe-se agora à premissa (4) do argumento. Ao examinar essa premissa, as diferenças entre o bayesianismo e a Estatística do Erro veem à luz de forma ainda mais evidente. Howson considera que, intuitivamente, a hipótese *J* (ausência de doença) deve ser considerada verdadeira, pois sua probabilidade posterior continua alta, não obstante o resultado *e* indicar que o indivíduo selecionado possui a doença. O cientista adepto da Estatística do Erro assume uma posição completamente diferente e acredita que *e* não é uma boa razão para afirmar a veracidade de *J*. Para a Estatística do Erro, a ausência da doença e, portanto, a veracidade de *J* apenas podem ser inferidas se todos os erros relativos a essas afirmativas tiverem sido excluídos. O indício *e*, por si só, para alguém adepto do enfoque de Mayo, não constitui um bom indício nem para considerar que *H* é verdadeira (doença presente), nem para realizar qualquer afirmativa relativa à ausência da doença. Seria necessário aprofundar a investigação e possivelmente, realizar novos testes, para que qualquer inferência confiável pudesse ser feita. Mayo indica que o problema com o enfoque bayesiano, nesse exemplo específico, é que ele sempre concluirá que não há doença, pois a doença é rara na população analisada e a probabilidade posterior de *J*, isto é, $P(J/e)$ sempre permanecerá alta. Esse resultado é perfeitamente coerente do ponto de vista bayesiano, mas não se coaduna com a Estatística do Erro, que está menos preocupada com a coerência de uma inferência do que com o fato de estar avalizada pela exclusão de erros possíveis.

Em suma, a tensão identificada entre bayesianismo e Estatística do Erro se expressa nos diferentes critérios apregoados por parte de cada enfoque para realizar inferências. De ponto de vista bayesiano, o que importa é conhecer a probabilidade posterior da hipótese em exame. Em alguma situações, a probabilidade prévia será idêntica à frequência e a probabilidade posterior poderá sustentar uma inferência eventualmente coincidente com a inferência que realizaria um adepto da Estatística do Erro. Em outras situações, porém, quando se está diante de um caso de probabilidade prévia como expressão do grau de crença do pesquisador na hipótese, as inferências avalizadas pelo bayesianismo e pela Estatística do Erro podem ser completamente diferentes.

Conforme o exemplo acima mostrou, a depender dos graus de probabilidade prévia assignados, uma hipótese pode ser inferida pelo bayesianismo sem que tenha passado por um teste severo segundo a concepção da Estatística do Erro. Para Howson, os resultados a que leva o bayesianismo são coerentes e são confirmados pelo senso comum. O fato de eles eventualmente autorizarem inferências incorretas se mostra como um problema menor, pois o acúmulo de indícios deverá corrigir isso no médio ou longo prazo. Para a Estatística do Erro, por outro lado, o objetivo primordial é realizar inferências tão avalizadas quanto o possível. Se não é possível ter certeza absoluta, deve-se, ao menos, controlar da melhor maneira possível todos os erros e incertezas envolvidas no caso.

CONCLUSÃO

Nesse capítulo foi brevemente apresentada a Estatística do Erro a partir dos escritos de um de seus autores seminais, Deborah Mayo. Trata-se de um enfoque que pretende se contrapor ao bayesianismo e que pouco tem em comum com ele, exceto o fato de buscar explicar o raciocínio científico levando em conta o conceito de probabilidade. Seus métodos e objetivos não poderiam contrastar mais com o bayesianismo. A lógica que rege a Estatística do Erro é o do rigor na análise de indícios para assegurar inferências tão seguras quanto o possível. De acordo com esse enfoque, a atribuição de probabilidades não se presta a expressar graus de confirmação ou de crença em hipóteses, mas para quantificar a capacidade de testes experimentais de diferenciar entre hipóteses alternativas, bem como permitir a detecção de erros de forma confiável. Não existe nenhum mecanismo padrão para a realização de inferências, equivalente ao Teorema de Bayes. Isso faz com que o enfoque que propõe a Estatística do Erro dependa muito mais de cada contexto experimental específico e que as inferências realizadas em uma situação particular sejam de difícil generalização. Se há algum elemento da Estatística do Erro que pode ser objeto de generalização, trata-se do conjunto de procedimentos que ensinam a identificar e evitar os erros experimentais mais comuns, ou seja, aquilo que Deborah Mayo chama de metaestasística.

Nada é mais distante do bayesianismo. Ao bayesianismo importa assegurar à indução uma lógica própria, capaz de desempenhar um papel equivalente ao da lógica clássica para o raciocínio dedutivo. Importa, ainda, explicar o raciocínio científico por meio de um algoritmo de aplicação direta, em que probabilidades expressam graus de confiança em determinada hipótese, antes e depois de ter sido contabilizado sobre essa hipótese o impacto confirmatório ou infirmativo de determinado indício. Essa abordagem sem dúvida aproxima mais a explicação do raciocínio científico do senso comum, mas também tem seus custos, como o risco de propiciar inferências incorretas, ainda que coerentes. Mas isso não é um problema maior para o bayesiano, que acredita que, já no médio prazo, o acúmulo de indícios levará o pesquisador na boa direção.

5. CONCLUSÃO

Ao longo desta dissertação, procurou-se apresentar o bayesianismo de Howson e Urbach como enfoque para analisar o raciocínio científico. Além de descrever como os indícios podem impactar as hipóteses segundo o Teorema de Bayes, seja para confirmá-las, seja para infirmá-las, viram-se também as soluções que esse enfoque propõe para questões conhecidas na Filosofia da Ciência, como o paradoxo dos corvos, as hipóteses *ad hoc* e o problema de Duhem, entre outros. A fonte primária utilizada para discorrer sobre o bayesianismo foi o livro já clássico de Howson e Urbach, *Scientific Reasoning – the Bayesian Approach* em suas diferentes edições.

Ao longo dos anos que separam a primeira (1989) e a última (2006) edição de *Scientific Reasoning*, Howson e Urbach responderam a várias críticas levantadas contra o bayesianismo, conforme vimos no capítulo 2. O empenho dos dois autores não foi capaz, porém, de sanar todas as objeções levantadas por autores inconformados com o excessivo grau de subjetivismo permitido pelo bayesianismo na análise do empreendimento científico, o que foi tratado no capítulo 3.

Determinados problemas foram identificados, para os quais o bayesianismo não parece ainda dispor de respostas satisfatórias. Segundo seus críticos, o enfoque bayesiano pode levar, por exemplo, ao descarte como *ad hoc* de hipóteses que são perfeitamente legítimas e explicam adequadamente determinados fenômenos, como no caso da hipótese de existência de Netuno para dar conta das anomalias identificadas na órbita de Urano. Como o exemplo de Neyman acerca da distribuição de larvas em um terreno mostrou, é difícil imaginar, ainda, de que maneira o bayesianismo poderia levar à formulação de uma hipótese *h´* em substituição a uma hipótese *h* que foi infirmada, a menos que *h´* já fizesse parte do conjunto de hipóteses mutuamente exclusivas inicialmente consideradas para explicar o fenômeno em exame. Por fim, é possível argumentar que o bayesianismo não funciona em situações em que o quadro teórico não é fixo, nem explica adequadamente fenômenos constituídos por uma sequência de eventos dependentes, como o jogo “vermelho ou azul”.

Outras questões ainda carecem de resposta bayesiana adequada. É o caso do chamado critério da simplicidade, ou seja, a conhecida preferência dos cientistas por hipóteses mais simples quando tem de escolher entre hipóteses rivais em igualdade de condições. Se a atribuição de probabilidades prévias é uma questão eminentemente subjetiva, como o bayesiano explica a atribuição de probabilidade prévia mais alta a uma hipótese mais simples em comparação com outra? Por fim, ao não se pautar por testes estatísticos rigorosos, o bayesianismo pode levar a que uma hipótese falsa seja mantida apenas porque sua probabilidade posterior não é suficientemente baixa. O bayesianismo não seria confiável, portanto, como método para identificar qual hipótese manter e qual hipótese descartar em face de uma instância de infirmação de uma teoria, ou seja, não daria conta do problema de Duhem.

São muitas as virtudes que exibe, não obstante, o bayesianismo. Por meio do Teorema de Bayes, é possível relacionar indícios e hipóteses de forma simples e direta. Ao se calcular como probabilidade posterior o impacto de confirmação ou infirmação de indícios sobre hipóteses, o cientista bayesiano tem à sua disposição um mecanismo que autoriza a realização de inferências científicas de maneira coincidente com o senso comum. Além disso, os princípios bayesianos permitem analisar de forma convincente várias questões inerentes ao raciocínio científico.

A difusão dos métodos bayesianos por áreas tão diversas como a filosofia da ciência, a teoria da decisão, a estatística e a análise de risco, entre outras, constitui uma mostra de que, a despeito da opinião de seus detratores, o bayesianismo não pode ser posto de lado sem maiores problemas. É forçoso reconhecer, porém, que, como toda e qualquer escola filosófica, o bayesianismo também apresenta limites. De fato, a aplicação do Teorema de Bayes, por si só, não garante a realização de inferências corretas tomadas isoladamente. Para o bayesiano, a veracidade ou falsidade de uma hipótese será declarada à luz da acumulação de indícios. O bayesiano acredita que, no médio prazo, os indícios indicarão o caminho correto. Mas esse é um risco que um adepto da estatística do erro não está disposto a correr.

No último capítulo dessa dissertação, apresentei, de forma resumida, o enfoque da estatística do erro conforme desenvolvido por Deborah Mayo em *Error and the Growth of Experimental Knowledge* e em alguns escritos dos anos 1990. Para a estatística do erro, o avanço da ciência e o raciocínio científico somente podem ser compreendidos em sua inteireza através da análise da atividade experimental. Nada interessa menos do que o grau de crença do pesquisador na hipótese em exame. Nada interessa mais do que investigar as margens de erro de determinado experimento e, a partir desse dado, assegurar que qualquer inferência será feita de forma tão segura quanto possível.

Há muito que se poderia dizer acerca da Estatística do Erro, mas isso estaria fora do escopo desta dissertação. O último livro de Deborah Mayo, *Error and Inference*, de

2010, reúne ensaios de autores consagrados e enfoca complexas questões filosóficas e metodológicas relativas à *error statistics* que ainda carecem de soluções definitivas. Discussões metodológicas, embora estejam fora do campo de ação bayesiano, constituem elemento essencial para a Filosofia da Ciência e autores como Popper, Carnap, Kuhn e Lakatos nunca chegaram, em seus escritos, a realmente investigar como questões do dia a dia dos laboratórios podem ser resolvidas para proporcionar conhecimento mais confiável. Esse é o caminho que a *error statistics* tem a oferecer.

Se alguma conclusão pode ser extraída dos capítulos precedentes é a de o bayesianismo se mantém como um enfoque frutífero para analisar o raciocínio científico. Isso não significa, em absoluto, que algumas das críticas que lhe são dirigidas possam ser descartadas sem maiores problemas, nem que enfoques alternativos como a *error statistics* não constituam contribuição de grande valor para a filosofia da ciência. De fato, o debate entre o bayesianismo e a estatística do erro parece estar longe de ter sido encerrado e certamente mereceria análise mais detalhada. A continuidade desse debate se dará menos em função da capacidade de um lado de desacreditar completamente o outro e mais em razão de que, na tentativa de analisar o raciocínio científico, ambas as propostas têm objetivos distintos e exibem considerável valor heurístico, cada uma a seu modo. Se esta dissertação tiver contribuído para tornar essas questões mais familiares à comunidade filosófica, terá cumprido sua função.

REFERÊNCIAS BIBLIOGRÁFICAS

CORFIELD, D. & WILLIAMSON, J. **Foundations of Bayesianism.** Dordrecht/Boston/London: Kluwer, 2001.

DE FINETTI, Bruno. Foresight: its logical laws, its subjective sources. In: Kyburg, H.E. e Smokler, H.E.(eds), **Studies in Subjective Probability,** Wiley, 1964.

__________________ .**Probability, Induction and Statistics**. New York: Wiley, 1972.

__________________ . Probabilism. In: **Erkenntnis**, 31, pp. 169-223, 1989.

DORLING, Jon. Bayesian Personalism, the Methodology of Research Programmes, and Duhem´s Problem. In: **Studies in History and Philosophy of Science,** volume 10, pp. 177-187, 1979.

EARMAN, John. **Bayes of Bust? – A Critical Examination of Bayesian Confirmation Theory.** Cambridge, MA: The MIT Press, 1992.

FEYERABEND, Paul. **Against Method (fourth edition)**. Verso, maio de 2010.

FORSTER, Malcom e SOBER, Eliott. Why Likelihood? In: TAPER, M. & LEE, S. (eds) **The Nature of Scientific Evidence**. Chicago: University of Chicago Press.

GIERE, Ronald. **Understanding Scientific Reasoning**. New York: Holt, Rinehart, 1984.

GILLIES, Donald. **Philosophical Theories of Probabiliy**. New York: Routledge, 2000.

_________________. Bayesianism versus Falsificationism. **Ratio**, 3, pp. 92-98, 1990.

_________________. Bayesianism and the Fixity of the Theoretical Framework. In: Corfield, D and Williamson, J, (eds.) **Foundations of Bayesianism,** pp 363 - 379. Kluwer: Dordrecht,Boston, London.

GLYMOUR, Clark. Why I Am not a Bayesian. In: CURD, M. & COVER, J. A. (eds) **Philosophy of Science – The Central Issues**. New York/London: W.W.Norton & Co. pp. 594-606.

__________________ . **Theory and Evidence.** Princeton: Princeton University Press, 1980.

GOOD, Irving. The Bayesian Influence, or how to sweep subjectivism under the carpet, in: HARPER & HOOKER (eds.) **Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science.** Vol. II, Dordrecht: Reidel, 1976.

HACKING, Ian. Slightly More Realistic Personal Probability. In: **Philosophy of Science**, volume 34, pp. 311-325, 1967.

HEMPEL, Carl. **Philosophy of Natural Science**. Englewood Cliffs: Prentice Hall, 1966.

______________. Studies in the Logic of Confirmation. In: **Mind,** volume 54, n. 213 (Janeiro de 1945), pp. 1-26.

HESSE, Mary. Bayesian Methods and the Initial Probabilities of Theories. In: FEIGL, H. & MAXWELL, G. (eds.) **Induction, Probability and Confirmation – Minnesota Studies in the Philosophy of Science.** Minneapolis: University of Minnesota Press, 1975.

HORWICH, Paul**.** Wittgensteinian Bayesianism. In: CURD, M. & COVER, J. A. (eds) **Philosophy of Science – The Central Issues.** New York/London: W.W.Norton & Co. pp. 607-624.

HOWSON, Colin & URBACH, Peter. **Scientific Reasoning: The Bayesian Approach.** Chicago: Open Court, 2006 (1993) (1989).

HOWSON, Colin. Error Probabilities in Error. In: **Philosophy of Science 64** (Junho de 1997).

________________ . **Hume´s Problem: Induction and the Justification of Belief**. Oxford: Clarendon, 2003 (2000).

________________ . **Against God**. Cambridge: Cambridge University Press, 2011.

JEFFREYS, Harold. **Theory of Probability**. Oxford: Clarendon, 1961 (1983).

JAYNES, E.T. Bayesian Methods: General Background. In **Maximum-Entropy and Bayesian Methods in Applied Statistics,** by J. H. Justice (ed.). Cambridge: Cambridge Univ. Press, 1986.

_____________ **Probability Theory: the Logic of Science**. Cambridge University Press, 2003.

KOLMOGOROV, Andrey. **Foundations of the Theory of Probability.** New York: Chelsea Publishing, 1950.

KUHN, Thomas. **The Structure of Scientific Revolutions.** Chicago: University of Chicago Press, 1970 (1962).

LAKATOS, Imre. Falsification and the Methodology of Scientific Research Programmes. In: **Criticism and the Growth of Knowledge.** Cambridge: Cambridge University Press, 1970.

MAHER, Patrick. Why Scientists Gather Evidence. In: **British Journal for the Philosophy of Science**, volume 41, pp. 103-119, 1990.

_______________ . Acceptance Without Belief. In: ***PSA 1990***, volume 1, pp. 381-392, 1990.

MAYO, Deborah. **Error and the Growth of Experimental Knowledge.** Chicago/London: The University of Chicago Press, 1996.

________________. Response to Laudan and Howson. In: **Philosophy of Science *64*** (Junho de 1997):323-333.

________________. Error Statistics and Learning from Error: Making a Virtue of Necessity. In: **Philosophy of Science, volume 64**, Suplemento. The University of Chicago Press, 1997.

________________. **Error and Inference**. Cambridge University Press, 2010.

MILLER, D. On the Maximization of Expected Utility. **PPE Lectures**, Lecture 8, Department of Economics: University of Vienna.

MUSGRAVE, Alan. Popper and Diminishing Returns from Repeated Tests. In: **Australasian Journal of Philosophy**, volume 69, pp. 124-134.

PEARSON, Egon Sharpe. On Questions Raised by the Combination of Tests Based on Discontinuous Distributions. In: **Biometrike** 37: 383-398, 1950.

PHILLIPS, Lawrence. **Bayesian Statistics for Social Scientists**. London: Nelson, 1973.

POPPER, Karl. The Propensity Interpretation of the Calculus of Probability and the Quantum Theory. In: KÖNER, S. **Observation and Interpretation,** Bristol: Butterworths Scientific Publications, 1957.

______________. The Propensity Interpretation of Probability. **British Journal for the Philosophy of Science,** 10 (37), 1959.

______________. **The Logic of Scientific Discovery.** London: Routledge, 2002 (1972) (1959).

______________ . **Conjectures and Refutations**. London: Routledge, 1963.

______________ . A Proof of the Impossibility of Inductive Probability. In: **Nature,** volume 302, pp. 687-88, 1983.

PORTUGAL, Agnaldo. Probabilidade e Raciocínio Científico. In: **Episteme,** Porto Alegre, v. 18, pp. 19-40, 2004.

REICHENBACH, Hans. **The Theory of Probability.** Berkeley: University of California Press, 1949.

SALMON, Wesley. **Foundations of Scientific Inference.** Pittsburgh: University of Pittsburgh Press, 1969.

________________. Rationality and Objectivity in Science or Tom Kuhn meets Tom Bayes. In: COVER J.A., CURD M., PINCOCK C (eds.). **Philosophy of Science: the central issues.** W.W. Norton Company, 1998.

SHAFER, Glenn. Constructive Probability. In: **Synthese**, volume 48 (1), pp. 1-60, 1981.

SOBER, Eliott. Bayesianism -- Its Scope and Limits. In R. Swinburne, (ed.), **Bayes' Theorem, Proceedings of the British Academy Press**, vol. 113, 2002, pp. 21-38.

STREVENS, Michael. The Bayesian Approach to the Philosophy of Science. In: **Macmillan Encyclopedia of Philosophy**. Macmillan Reference, 2005.

SWINBURNE, Richard. **Epistemic Justification.** Oxford Univeristy Press, 2001.

Teller, P. (1973), Conditionalization and Observation. **Synthese** *26:* 218-258.